Türk Üstün Zekâ ve Eğitim Dergisi
2013, Cilt 3, Sayı 2, 66-80

Turkish Journal of Giftedness and Education
2013, Volume 3, Issue 2, 66-80

Revalidating the Arabic Scale for Teachers' Ratings of Basic Education Gifted Students' Characteristics Using Rasch Modeling

Arapça Temel Eğitim Üstün Zekâlı Öğrencilerin Özellikleri Öğretmen Değerlendirme Ölçeğinin Rasch Modelleme ile Yeni Geçerliği

Salah Eldin Farah Atallah Bakheit[1]

Abstract

The Arabic scale for teachers' ratings of basic education gifted students' characteristics is one of the most common Arabic measures used for initial identification of gifted students in some Arabic countries. One of the shortcomings of this scale is that it is based on the classical theory of measurement. This study sought to revalidate the scale in the light of Rasch modeling which rests upon the modern theory of measurement and to develop different criteria for interpreting the levels of individuals' traits. The scale was administered to 830 of Basic Education students in Khartoum (ages ranged from 7 to 12 years). Two groups of students participated in the study: a calibration sample (N = 250) and a standardization sample (N = 580). The statistical treatments were performed using the PSAW 18 and RUMM 2020 programs according to Rasch's unidimentional model. Six of the scale's items were deleted for not conforming to Rasch Modeling. This left the scale with 31 items. Besides, new criteria for the scale were developed by obtaining the t-scores and special education scores that match the various ratings of the individuals' ability.

Key Words: Arabic Scale for Teachers' Ratings of Basic Education Gifted Students' Characteristics, gifted students, Rasch Modelling, revalidation

Öz

Arapça Temel Eğitimde Üstün Zekâlı Öğrencilerin Özellikleri Öğretmen Dereceleme Ölçeği, bazı Arap ülkelerinde üstün zekalı öğrencilerin tanılanmasında en sık kullanılan Arapça ölçeklerden biridir. Bu ölçeğin eksikliklerinden biri klasik ölçme kuramına dayanmasıdır. Bu araştırmanın amacı modern ölçme kuramına dayanan Rasch Modeli ile ölçeğin geçerliğini yeniden incelenmek ve bireysel özelliklerin düzeylerini yorumlamada yeni kriterler geliştirmektir. Ölçek Kartum'da 830 temel eğitim öğrencisine (7-12 yaş aralığı) uygulanmıştır. Katılımcılar ayarlama (kalibrasyon) örneklemi (N=250) ve standardizasyon örneklemi (N= 580) olmak üzere iki guruba ayrılmıştır. İstatistiksel uygulamalar tek boyutlu Rasch modeline göre PSAW 18 ve RUMM 2020 programları kullanılarak yapılmıştır. Rasch modeline uymadığı için ölçekteki altı madde çıkartılmış ve geriye 31 madde kalmıştır. Bunula birlikte çeşitli bireysel yeteneklerin derecelemelerine uyan özel eğitim ve t-puanları elde edilerek ölçek için yeni kriterler geliştirilmiştir.

Anahtar Sözcükler: Arapça Temel Eğitimde Üstün Zekâlı Öğrencilerin Özellikleri Öğretmen Dereceleme Ölçeği, geçerlik, Rasch modeli

Introduction

The Arabic scale for teachers' ratings of basic education gifted students' characteristics appeared in 1996 when it was published by the Arab League Educational Cultural and Scientific Organization (ALECSO) in its pioneering guide entitled "A Guide for the Identification

[1] *PhD, Associate Professor, Special Education Department, College of Education, King Saud University, Riyadh, Saudi Arabia; bakhiet@ksu.edu.sa*

ISSN 2146-3832, http://www.tuzed.org

of Gifted Students in Basic Education." This scale, unlike other tools that were derived from foreign tests and adapted to suit the Arabic environments, was originally designed for the Arabic environment. The scale was developed by a group of education and psychology experts (see Abdulghafar, et al, 1996) to rate the cognitive, motivational, affective and social characteristics of gifted students. The scale, which included 37 of gifted students' characteristics, was developed in the light of the characteristics that empirically proved to be relevant to mental superiority and literature on personal characteristics of giftedness. Because of the multiplicity of the situations in which the teacher observes students closely, teacher's ratings are likely to tap the basic patterns of student characteristics that relate to creative behavior. Creativity relates to curiosity, imaginativeness, preference for multiple solutions, discovery of new relations, and the ability to express ideas clearly (Shaker, 1995).

The scale serves two purposes. First, it allows for initial identification of the characteristics of gifted students. Second, it directs teachers' attention to types of characteristics that normally go unnoticed. Added to this, it also helps the determination of students' eligibility for gifted programs.

The scale is one of the important tools developed in the Arabic environment. It was piloted and used in four Arabic countries: Egypt (Sadeq, Assayed & Alaam, 1996), Tunisia (Ben Fatemah, Moawyiah & Asweisi, 1996), United Arab Emirates (Atahan & Abo-Hilal, 1996) and Iraq (Alhamadani, Rasoul & Aleigili, 1996). It was also approved as a basic tool for initial identification of gifted students in Alqabas schools in Khartoum. Besides, it was adopted in the project of the Sudanese Ministry of Education for the identification of gifted students.

The research conducted by ALECSO in Egypt (Sadeq, Assayed & Alaam, 1996), Tunisia (Ben Fatemah, Moawyiah & Asweisi, 1996), United Arab Emirates (Atahan & Abo-Hilal, 1996) and Iraq (Alhamadani, Rasoul & Aleigili, 1996) revealed that the scale is characterized with:

1. A high ability for identifying and predicting gifted students in regard to general achievement both in the sixth and ninth grades.
2. An average ability for identifying giftedness in arts, the Arabic language, science and mathematics among the sixth grade students, and in arts, the Arabic language and mathematics among the ninth grade students.
3. A weak ability for identifying giftedness in the Arabic language among the ninth grade students.

The scale had an alpha reliability coefficient of .93 in the study conducted in United Arab Emirates by Atahan and Abo-Hilal (1996). However, it is noteworthy that the scale was not subjected to factorial analysis, and thus the dimensions of the scale were not identified and it was considered a unilateral measure, even though its developers allege that it assesses characteristics in four dimensions.

Bakheit (2006) applied the scale to students in the second cycle of basic education in Alqabas schools in Khartoum state. The researcher piloted the scale on a sample of 58 students and then validated it on a sample of 955 students (52.9% males and 47.1% females). The scale yielded good indices of face validity and internal consistency. The reliability coefficients for the items of the whole scale ranged between .31 and 84. The reliability coefficients if item deleted ranged between .38 and .81. All coefficients were significant at the .01 level. The means of the items ranged between 1.5 and 2.5 and standard deviations ranged between .53 and .84. The discrimination index (between top and bottom quartiles) was computed (t = 18.14, p = .01). The scale also achieved reasonable concurrent validity as it correlated with Renzulli's measures of behavioral characteristics (r = .60). It yielded an authentic validity coefficient of .61 (p = .01). The similarity percentage was 77.1%. The values of experimental validity by correlation between the scale and intelligence (estimated by measure of the standard progressive matrices), mathematics, achievement, and creativity were .29, .51, .58, .17 (all significant at the .01 level) respectively.

The reliability of the scale was established by several techniques. It yielded an alpha coefficient (for internal consistency) of .76. Alpha coefficients if item deleted ranged between .75 and .76. The split half technique (for the correlation between odd and even items) gave a value of .97. After modification, the scale yielded reliability values of .98 and .79 by Spearman-Brown and Guttmann respectively. The test retest techniques gave a reliability value of .83. The local criteria of grades (grade level) and gender were elicited, and the used criteria were t-scores.

Another examination of the scale was conducted in the Sudan by Bakheit (2008) who sought to investigate the psychometric characteristics of the scale in the Sudanese environment and to elicit criteria for it. The researcher administered the scale to 2216 of basic education students (57.5% males and 42.5% females) whose ages ranged between 8 and 12 years (M = 9.8 and SD = 1.2). The validity of the scale was then established using content validity, factorial analysis and hypothetical construction validity. The results revealed good validity indices. The referees concurred that the scale is valid without any modification and factor analysis proved that the scale has five factors. The scale items correlated significantly at the .01 level with its dimensions and with the whole sale. The reliability of the scale was established using internal consistency (alpha coefficients ranged between .76 and .97) and split half. After modification, the reliability was reestablished using Guttmann's and Spearman-Brown's formulas (reliability coefficients ranged between .92 and .5). Besides, percentiles for each gender and for the total sample were computed. Most percentiles revealed that scores are graded according to chronological age.

The development of several observation checklists and scales that are used by teachers and parents, help us to develop information tables concerning students' strengths such as the ability to generate creative solutions to problems and their level of motivation (Elliott, Busse, and Gresham, 1993; Feldhusen and Heller, 1986; Renzulli , Siegle, Reis, Gavin, & Reed, 2009).

Based on a survey of relevant literature, Rust (1985) and Brody (2007) highlighted the importance of such tools for the identification of gifted students.

The behavioral trait approach established a link between the psychometric perspective and the behavioral impressionistic perspective which is based on observation in the identification of gifted individuals. This approach has therefore received a widespread interest on the part of researchers and theorists. Because of the importance of scales for rating the characteristics of gifted children, scales have been widely used in the identification of gifted individuals. The literature on giftedness is rich in identification tools (Pfeiffer & Jarosewich, 2007; Pfeiffer, Petscher, & Jarosewich, 2007; Pfeiffer & Petscher, 2008; Pfeiffer, Petscher, & Kumtepe, 2008; Renzulli , Siegle, Reis, Gavin, & Reed, 2009) .

Despite the importance of all local, regional and international scales, measures and tests for identifying giftedness and their contribution to address bias in the identification of gifted children in minorities (hence, solving the problem of their underrepresentation in programs for gifted individuals), all these tools were developed in the light of the classical (traditional) theory in psychological measurement, which has many shortcomings. Doubts are therefore raised concerning the psychometric characteristics of such tools and the accuracy and objectivity of their use in measuring human behavior. Researchers and experts summarized the shortcomings of the classical theory as follows: The total test score being restricted to test items; lack of linearity of measurement; measurement in more than one dimension; uniformity of test scores and the level of the measured variable; and the variability of the meaning of test items across time (Adardeer, 2004). A fixed measurement unit is lacking since measurement locations are not placed on the variable continuum linearly. The dependence of individuals' scores on test items may result in variability in the distance between every two consecutive scores. This results in the variability of the quantitative meaning of any specified difference across the range of test scores (Kathem, 1996). Besides the characteristics of test items, e.g. item difficulty and item discrimination are affected by the individuals' ability. The same item is easy for students with high level abilities and difficult for those with low level abilities. And if the sample is comparatively homogenous, the values of discrimination coefficients are smaller than those obtained from a heterogeneous sample (Hambleton &Swaminathan, 1989). An individual's score on a given test is affected by test items. The individual obtains higher scores when tested with easy items than with difficult ones. This way the individual's true ability is not assessed accurately. Thus, the result of measurement varies from test to test (Alaam, 2000); Comparison between individuals in the trait or ability assessed by the test entails applying the same test items or equivalent group of items to each individual. Hence, we cannot compare levels of ability if individuals answer items that are different in difficulty (Abdelmaseih, 1991). Test reliability is affected by the testing situation. Test reliability according to this theory is established either by administering the test twice to the sample (test re-test method) or by using two equivalent versions of the test. However, the testing situation can be different in the two applications and it is rather difficult to develop

two equivalent versions of the same test, which affects test reliability (Hambleton & Swaminathan, 1989). The variance of measurement errors for the entire sample can be equal though the performance of some individuals can be more consistent than the performance of others and though the degree of this consistency varies according to the level of the individuals' ability or the level of the ability measured by the test (Randall, 1998: 6, cited in Abo-Hashem, 2006);This theory does not present a psychological interpretation of how the individual tries to answer a test item. This interpretation is important if we wish to predict the characteristics of the scores derived from a given population or different populations, or if we wish to design tests with given psychometric characteristics that suit a given population. In addition, the meaning of test items changes with time. Environmental and testing circumstances are always subject to changes. The deletion or change of any of the test items can change the subjects' scores. This change is difficult to predict (Alaam, 1985); All the characteristics of the tests that are based in its construction on the classical theory such as difficulty, discrimination and reliability coefficients depend on the characteristics of the sample taking the test and the characteristics of the items in the test (Alaam, 1987).

For the previously mentioned shortcomings and with the advent of the modern theory in psychological measurement (what is called the latent trait theory or item response theory), test developers began to develop tools in the light of the modern theory that proved to be highly important in developing the psychometric characteristics of psychological and educational tests.

Those who are knowledgeable of research in the field of giftedness and superiority may notice that measurement models of the modern theory are seldom used. The researcher found, by surveying published research, that there are only five studies up to 2010 employing modern models of measurement for the identification of gifted individuals. The use of such measures in the five studies is restricted to cognition and achievement. No scales or checklists identifying the behavioral characteristics of superior individuals were used in the five studies.

The present study thus seeks to meet a basic need in the field by developing a scale for teacher's ratings of basic education gifted students' characteristics according to Rasch's unilateral model. This is expected to provide a suitable tool for the initial identification of gifted individuals.

Problem

The identification of giftedness at the national and international levels used to depend on tools developed in the light of the classical theory in psychological measurement whether in initial identification or final assessment. This had negative effects on the field since the accuracy and psychometric characteristics of such tools are questioned. It is therefore crucial that tools used for the identification of gifted individuals be developed in the light of modern models to achieve more accurate measurement. The widespread use of the Arabic scale for

teacher's ratings of basic education gifted students' characteristics in Arabic countries also urged the researcher to conduct the present study to subject the scale to Rasch Modeling in an attempt to answer the following questions:

1. To what extent does the scale for rating basic education gifted students' characteristics conform to Rasch Modeling?
2. What is the rating of individuals' trait for every possible total score on the scale?
3. How reliable is the scale after calibrated using Rasch Modeling?
4. Does the scale have acceptable criteria?

Method

Participants

Participants in the present study were divided into two groups: the calibration sample used for calibrating test items and the standardization sample used for obtaining the scale's criteria.

The calibration sample includes 250 of basic education students (125 males and 125 females) in Khartoum whose ages ranged from 7 to 12 years, i.e. students from the second to the sixth grade. The standardization sample consisted of 580 students (290 males and 290 females) whose ages ranged from 7 to 12 years, students were from the second to the sixth grade, 20% for each grade.

The Instrument

The researcher used the Scale for Teacher's Ratings of Basic Education Gifted students' Characteristics, which is approved by LECSO. This scale consists of 37 items, each representing a trait of gifted students. The items cover the cognitive, motivational, affective and social aspects of giftedness (Appendix 3) (Alhamadani, Rasoul, and Alejeili, 1996; Abdulghafar, Sadeq, Assyed, Beshara, Atahan and Alaam,1996; Atahan & Abohilal, 1996; Bakheit, 2006, 2008; Sadeq, Alboni, Besharah, Abohatab, Rabie and Ben Fatema, 1996; Sadeq, Assyed, and Alaam, 1996).

Instructions for the completion of the scale: The scale seeks to view the classroom teacher's opinion concerning the degree to which each of the characteristics included in the scale applies to each student. The teacher provides the data about the student and then puts a tick in the cell that represents, in his point of view, the degree to which the characteristic included in the item applies to the student. A three point likert-scale is used: high (= 2 marks), average (= 1) and low (= 0).

Procedures

The test was administered in the first semester of the school year 2010/2011. Teachers were given copies of the scale to rate their students' characteristics included in the items. The

completion of the scale was not timed and the teachers were asked not to leave any items unanswered. Answers were coded and fed to the PSAW 18 statistical program. Data were treated and fed to the RUMM 2020 program to analyze it using Rasch's unilateral model to calibrate items. The RUMM 2020 program deals with scores 0, 1 and 2, which is not possible with other statistical programs using Rasch Modeling or even the old version of the RUMM program. Criteria interpreting individuals' different levels were computed. T-scores and special education scores matching ability ratings (special education scores are computed by multiplying the standard score by 15 plus 1000) were computed.

Results

Items of the scale were analyzed to investigate their internal consistency and to make sure that each item measures the intended characteristic. Using Rasch Modeling, the researcher analyzed responses to the scale's items through the RUMM 2020 program.

Eliminating complete and zero data from the analysis matrix. This included: Eliminating individuals who obtained the full mark. The ability of such individuals is higher than the range covered by the scale. This resulted in eliminating 23 individuals of the calibration sample. Analysis was then performed on the remaining 227 individuals. Eliminating individuals who failed to obtain any score. The ability of such individuals is lower than the range covered by the scale. No individuals were eliminated on this basis since all individuals obtained scores. Eliminating any item whose characteristic is met by all students. This led to the elimination of the second item, leaving the scale with 36 items. Eliminating any item whose characteristic is not met by any student. Such items cannot discriminate between levels of the variable. No item was deleted based on this procedure.

Eliminating the individuals who are not appropriate to the model. After completing the previous step, analysis was performed to eliminate the individuals who are not appropriate for the model, i.e. individuals who are not appropriate for the calibration process according to the following criteria: Eliminating the individuals whose appropriateness values are less than -2. This means that the rating obtained by such individuals is similar, which means that responses are not valid. Eliminating the individuals whose appropriateness values are more than +2. This means that such individuals exceeded the statistically acceptable limit by having the characteristics of the items whose characteristics are higher than theirs, or lacking the characteristics of the items whose characteristics are lower than theirs. This means that raters did not rate these characteristics accurately. The previous two steps led to the elimination of 25 students who were either higher or lower than the appropriateness limits. This way only students with valid responses were kept (N = 202).

Eliminating the items that are not appropriate for the model. Data were re-analyzed to eliminate the items that are not appropriate for the model, i.e. items that have defects which make them inappropriate to calibrate the measured variable. This was done according to the following criteria: Eliminating the items whose appropriateness values are less than -2.5 since

this means that such items are not independent of the other items in the scale or that they measure a variable which is highly similar to the measured one. Eliminating the items whose appropriateness values are higher than +2.5 since this means that there is a defect in the construction of the item or that it measures another variable. Based on this analysis, 5 items were deleted from the scale (these included items: 9, 10, 25, 27 and 28). This procedure left the scale with 31 items. After performing the procedures in steps 1, 2 and 3, the researcher reached the final calibration of the scale. This is listed in table 1 (in Appendix 1).

Validity of the Scale

Validity of Calibration. The calibration of a group of items measuring the same characteristic on a common shared scale using Rasch Modeling means that such items meet the condition of the unilateralism of measurement (i.e. they define one variable). The unilateralism of measurement that Rasch Modeling secures establishes the validity of the calibration of items in measuring the target variable. It also achieves the validity of the calibration of individuals' abilities on the variable continuum, which is based on the validity of their responses to the items (Kathem, 2000).

The unilateralism of measurement is achieved through the appropriateness of all individuals and items to the model according to the criteria used in analysis. These criteria show the extent to which the appropriateness of the item program (RUMM 2020) expresses what the remaining items express on the continuum of the target variable. It also shows the extent to which the use of Rasch Modeling is consistent in developing the scale for rating basic education gifted students' characteristics.

Factorial analysis. The researcher conducted factorial analysis on the calibration sample using the maximum likelihood procedure. The KMO equivalence coefficient was .88. Bartlett's test of spherity and Ki Square was 2408.557. The degree of freedom was 136 (significant at .0001. The correlation coefficients between items were high (ranging from .46 to .74). The most important result was that the items loaded on one factor explaining 61% of variance. The appropriateness quality of the scale (Ki Square) was 164.57, with degrees of freedom of 61 and significance level of .0001. All these values show that the scale is valid.

Reliability of the Scale

Reliability of Calibration. The calibration of items on a common calibration scale according to Rasch Modeling after the deletion of all the items that are not appropriate for measurement and the individuals who are not appropriate for the model means that the conditions of the model are met (e.g. the independence of measurement). It also means that difficulty and ability ratings are reliable, and that they are not affected by the diversity of the group of items taken from the original calibration scale or the diversity of the individuals who took the test.

The reliability coefficient. The RUMM 2020 establishes reliability according to the classical theory of measurement. The alpha coefficient for the scale's reliability was .97.

Computation of the Test Criteria

The role of Rasch Modeling is limited to calibrating the scale's items according to their correspondence with the model and rating the levels of the individuals on the scale with the units of logit and minf. The computation of test criteria was performed using group referenced criteria that match ability ratings on the test. The criteria were computed in the following steps:

- Computing the total raw score for every participant in the standardization sample (580) on the final form of the scale (31 items) after deleting inappropriate items.
- Converting the total raw scores for all the participants to the matching ability ratings using tables for rating the matching probable ability for every probable total score on the scale.
- Computing the mean and standard deviation for the individuals' ability on the scale estimated by the minf unit by converting logit to minf with the linear transformation equation.
- Computing the criteria of t-scores and special education scores for each ability rating estimated by the minf unit. (in table 2, appendix 2)

Discussion

The study aimed to recalibrate the Teacher Rating Scale of the Characteristics of Gifted Children in Basic Education Phase for rating basic education gifted students' characteristics using Rasch's unilateral model and to develop different criteria for interpreting individuals' ability levels. Six items were eliminated from the scale because of their inappropriateness to Rasch Modeling. Hence, the scale had 31 items after recalibrated. The study also established criteria for the scale by obtaining t-scores and special education scores matching the various ratings of individuals' ability.

The study benefited from the linearity of measurement that characterizes Rasch Modeling since there is one measurement unit for item difficulty and the individual's characteristic, the logit unit which is converted in the present study to the minf unit, special education scores and t-scores. The number of the deleted items according to the convergent appropriateness criterion is comparatively small (6 items) compared with the findings of studies on personality measurement in general. The study also revealed that there is a difference in the order of the scale items before and after calibration using Rasch Modeling and that the order after calibration is more logical than the order before calibration.

The results of the study direct the attention of gifted programs specialists to the significance of latent trait models in analyzing the data for the identification of gifted individuals using the latent trait theory. They also stress the need to base the identification process not only on raw scores but also on the results of the latent trait theory and modern psychological measurement models to obtain accurate results. This, in turn, results in making right decisions

about gifted students. This way the study provided the Arabic environment with an objective and accurate scale that has definite calibration. This scale can be used by individuals in charge with gifted programs, researchers and teachers for identification and research purposes without having to pilot it to establish its psychometric characteristics since the calibration of individuals is not affected by items, and the calibration of items is not affected by individuals.

The study reflected the value of using Rasch Modeling in developing measures for identifying gifted individuals in general and measures of gifted individuals' characteristics in particular. The researcher therefore offers the following recommendations: Using the model in developing more measures of giftedness and superiority to eliminate the shortcomings that used to be raised about these measures. Using the three-dimension model to analyze the items of the scale for rating basic education gifted students' characteristics to study the effect of guessing and the discrimination power of the items.

Funding

The author is thankful to the Deanship of Scientific Research, College of Education Research Center at King Saud University for funding this research.

References

Abdelmaseih, E. (1991). Using Rasch's logarithmic one-parameter model in analyzing the items of criterion referenced bipolar cognitive tests: An experimental study. *The Journal of Research in Education and Psychology, 4*, 443-475.

Abdulghafar, A., Sadeq, A., Assyed, A., Beshara, G., Atahan, K., & Alaam, S. E. (1996). The scale for teacher rating basic education gifted students. In Sadeq et al (Eds.), *A guide for the identification of gifted students in basic education* (pp. 95-97). Tunisia: LECSO.

Abo-Hashem,A.(2006). A comparative study between the traditional theory and Rasch modeling selection in paragraphs entrances scale study of students at the university. *Journal of the College of Education at the University of Zagazig*, 52, 17-70.

Adardeer, A. (2004). *Contemporary studies in educational psychology.* Cairo: Alam Alkotob.

Alaam, S. E. (1985). Analyzing the data of mental tests using Rasch's logarithmic probable model: An experimental study. *The Arabic Journal for Human Sciences, 17*, 100-124.

Alaam, S. E. (1987). A comparative critical study of latent trait models and classical models in psychological and educational measurement. *The Arabic Journal for Human Sciences, 27*, 18-44.

Alaam, S. E. (2000). *Educational and psychological measurement and evaluation: Its basics, applications and modern trends*. Cairo: Dar Alfekr Alarabi.

Alhamadani, M., Rasoul, K., & Alejeili, S. (1996). The findings of field study in Iraq. In Sadeq et al (Eds.), *A Guide for the Identification of gifted students in basic education* (pp. 60-63). Tunisia: LECSO.

Atahan, M. and Abohilal, M. (1996). The findings of field study in United Arab Emirates. In Sadeq et al (Eds.), *A guide for the Identification of gifted students in basic education* (pp. 30-38). Tunisia: LECSO.

Bakheit, S. E. (2006). Identifying gifted students in the Sudan in the light of the a guide for the identification of gifted students developed by LECSO: Reliability and validity indices and local criteria. *The Arabic Journal of Education, 26*(1), 71-101.

Bakheit, S. E. (2008). The Sudanese version of the scale for rating basic education gifted students' characteristics "The scale of ALECSO for initial identification of gifted individuals. *The Journal of the Mission of Education and Psychology, 31*, 123-162.

Ben Fatema, M., Moawyiah, A., & Asweaisy, N. (1996). The findings of field study in Tunisia. In Sadeq et al (Eds.), *A guide for the identification of gifted students in basic education* (pp. 39-59). Tunisia: LECSO.

Brody, L. E. (2007). Review of the gifted and talented evaluation scales. In K. F. Geisinger, R. A. Spies, J. F. Carlson, & B. S. Plake (Eds.), *The seventeenth mental measurements yearbook* (pp. 343–345). Lincoln, NE: Buros Institute of Mental Measurement of the University of Nebraska—Lincoln.

Elliott, S. N., Busse, R. T., & Gresham, F. M. (1993). Behavior rating scales: Issues of use and development. *School Psychology Review, 22*, 313–321.

Feldhusen, J. F., & Heller, K. A. (1986). Introduction. In K. S. Heller & J. F. Feldhusen (Eds.), *Identifying and nurturing the gifted: An international perspective* (pp. 19–31). Toronto: Hans Huber.

Hambleton, R., & Swaminathan, H. (1989). *Item response theory principles and applications*. Boston: Kluwer Nijhoff Publishing.

Kathem, A. (1996). Latent trait models. In A. Asharqawy, S. Asheikh, A. Kathem & Nadia Abdesalam (Eds.), *Modern trends in psychological and educational measurement and evaluation (*pp. 281-430*)*. Cairo: Alanjelo Almasriah.

Kathem, A. (2000). *Modern trends in question banks in the educational basics of university teacher preparation*, 3rd edition. Cairo, Ain Shams University.

Pfeiffer, S., Jarosewich, T.(2007). The gifted rating scales-school form: an analysis of the standardization sample based on age, gender, race, and diagnostic efficiency. *Gifted Child Quarterly, 51(*1*),* 39-50.

Pfeiffer, S., Petscher, Y., & Jarosewich, T. (2007). The gifted rating scale preschool/kindergarten form: An analysis of the standardization sample based on age, gender, and race. *Roeper Review, 29(*3*),* 206-211.

Pfeiffer, S., Petscher, .(2008). Identifying young gifted children using the gifted rating scales preschool/kindergarten form. *Gifted Child Quarterly, 52(*1*)* ,19-29.

Pfeiffer, S ., Petscher, Y ., & Kumtepe, A. (2008). The gifted rating scales-school form: A validation study based on age, gender, and race. *Roeper Review, 30(*2*),* 140-146.

Randall, S.(1998). Comparing measurement theories. *Paper Present at the Annual Meeting of the American Educational Research Association* (San Diego, CA, April 13-17).

Renzulli, J. Siegle, D., Reis, S., & Gavin, M. (2009). An investigation of the reliability and factor structure of four new scales for rating the behavioral characteristics of superior students. *Journal of Advanced Academics, 21(*1*)*, 84-108.

Rust, J.(1985). Review of Scales for Rating the Behavioral Characteristics of Superior Students. In J. V. Mitchell Jr. (Ed.), T*he ninth mental measurements yearbook* (Vol. 2, pp. 312–313*).* Lincoln, NE: Burros Institute of Mental Measurement of the University of Nebraska—Lincoln.

Sadeq, A., Alboni, A., Besharah, G., Abohatab, F., Rabie, M., & Ben Fatema, M. (1996). *A guide for the Identification of gifted students in basic education*. Tunisia, LECSO.

Sadeq, A., Assyed, A., & Alaam, S.E. (1996). The findings of field study in Egypt. In Sadeq et al (Eds.), *A guide for the Identification of gifted students in basic education* (64-81). Tunisia: LECSO.

Shaker,S. (1995). *Psychology of creativity*. Cairo, Dar Ghareeb.

Appendix 1

Table 1: The final calibration of the scale for rating basic education gifted students' characteristics

Previous no. of the item	Logit	Minf	Standard error	Appropriateness of residuals	X2	Probability
1	-0.433	47.84	0.149	-1.477	6.85	0.032552
3	-0.289	46.77	0.134	-0.86	0.737	0.691931
4	0.029	48.56	0.135	-0.087	0.676	0.713078
5	0.104	50.15	0.133	-0.7	1.988	0.370072
6	-0.766	50.52	0.14	0.484	0.907	0.635473
7	-0.373	46.17	0.143	-0.813	3.119	0.210225
8	-0.442	48.14	0.137	-2.483	4.276	0.117901
11	-0.1	47.79	0.136	-0.525	4.403	0.110615
12	0.607	51.31	0.127	-0.487	0.258	0.878773
13	0.102	52.99	0.131	0.239	7.115	0.028508
14	0.575	49.5	0.126	-0.334	1.468	0.479931
15	-0.025	53.04	0.131	-1.901	4.642	0.098195
16	-0.191	50.51	0.138	-2.004	2.752	0.252622
17	0.226	52.88	0.136	-0.747	1.273	0.529087
18	-0.338	49.88	0.137	-0.366	0.879	0.644273
19	-0.803	49.05	0.139	0.709	4.583	0.101113
20	0.078	51.13	0.129	-0.643	0.081	0.960289
21	0.996	48.31	0.128	1.654	7.928	0.018989
22	-0.901	45.99	0.141	-0.311	0.836	0.65842
23	-0.464	50.39	0.134	-0.605	0.916	0.632639
24	0.688	54.98	0.131	1.261	3.473	0.176177
26	-0.013	45.5	0.133	2.172	14.906	0.000581
29	-0.332	47.68	0.133	-0.034	0.573	0.751001
30	0.494	53.44	0.126	-0.835	0.378	0.827821
31	-0.499	54.73	0.138	-0.632	0.906	0.635754
32	0.101	49.94	0.132	-2.385	4.673	0.096643
33	0.514	49.6	0.127	-1.204	1.294	0.523509
34	-0.76	50.28	0.138	-1.13	1.183	0.553582
35	0.307	48.34	0.131	-1.034	1.639	0.440543
36	0.45	52.47	0.132	-1.668	1.624	0.443976
37	0.325	47.51	0.128	-2.366	2.612	0.270905

Degrees of freedom 1 and 2 are 214.03 and 2 respectively

Appendix 2

Table 2: The Corresponding Ratings For Every Probable Total Raw Score On The Scale For Rating Basic Education Gifted Students' Characteristics

Corresponding ability					Corresponding ability				
Total raw score	With logit	With minf	t-score	Special education score	Total raw score	With logit	With minf	t-score	Special education score
0	**-**	**-**	**-**	**-**	**32**	-0.114	49.43	48.86	98.29
1	-4.769	26.16	2.31	28.47	**33**	0.226	51.13	52.26	103.39
2	-4.220	28.9	7.8	36.7	**34**	0.338	51.69	53.38	105.07
3	-3.845	30.78	11.55	42.33	**35**	0.449	52.25	54.49	106.74
4	-3.553	32.24	14.47	46.71	**36**	0.561	52.81	55.61	108.42
5	-3.311	33.45	16.89	50.34	**37**	0.672	53.36	56.72	110.08
6	-3.102	34.49	18.98	53.47	**38**	0.782	53.91	57.82	111.73
7	-2.916	35.42	20.84	56.26	**39**	0.893	54.47	58.93	113.4
8	-2.747	36.27	22.53	58.8	**40**	1.004	55.02	60.04	115.06
9	-2.590	37.05	24.1	61.15	**41**	1.115	55.58	61.15	116.73
10	-2.443	37.78	25.57	63.36	**42**	1.226	56.13	62.26	118.39
11	-2.304	38.48	26.96	65.44	**43**	1.338	56.69	63.38	120.07
12	-2.172	39.14	28.28	67.42	**44**	1.450	57.25	64.5	121.75
13	-2.044	39.78	29.56	69.34	**45**	1.564	57.82	65.64	123.46
14	-1.920	40.4	30.8	71.2	**46**	1.680	58.4	66.8	125.2
15	-1.810	40.95	31.9	72.85	**47**	1.798	58.99	67.98	126.97
16	-1.682	41.59	33.18	74.77	**48**	1.918	59.59	69.18	128.77
17	-1.566	42.17	34.34	76.51	**49**	2.041	60.21	70.41	130.62
18	-1.338	43.31	36.62	79.93	**50**	2.169	60.85	71.69	132.54
19	-1.298	43.51	37.02	80.53	**51**	2.302	61.51	73.02	134.53
20	-1.226	43.87	37.74	81.61	**52**	2.441	62.21	74.41	136.62
21	-1.114	44.43	38.86	83.29	**53**	2.588	62.94	75.88	138.82
22	-1.003	44.99	39.97	84.96	**54**	2.745	63.73	77.45	141.18
23	-0.892	45.54	41.08	86.62	**55**	2.914	64.57	79.14	143.71
24	-0.781	46.1	42.19	88.29	**56**	3.101	65.51	81.01	146.51
25	-0.669	46.66	43.31	89.97	**57**	3.311	66.56	83.11	149.67
26	-0.558	47.21	44.42	91.63	**58**	3.553	67.77	85.53	153.3
27	-0.446	47.77	45.54	93.31	**59**	3.845	69.22	88.45	157.68
28	-0.334	48.33	46.66	94.99	**60**	4.221	71.11	92.21	163.32
29	-0.222	48.89	47.78	96.67	**61**	4.771	73.86	97.71	171.57
30	-0.110	49.45	48.9	98.35	**62**	**-**	**-**	**-**	**-**
31	-0.002	49.99	49.98	99.97					

Appendix 3

The Scale for Rating Basic Education Gifted Students' Characteristics

Basic data:

Residence: ……. Date of Application: ……...

Student name: …… Birth date: ……..

Name of School: …… Student's age: …….

Grade: ……..

Name of the teacher who completed the scale: ……...

Teacher's Specialization: ……

No.	Statement	High	Average	Low
1	Possesses language vocabulary that exceeds his chronological age			
2	Expresses himself and his ideas clearly			
3	Possesses varied repertoire of knowledge			
4	Discovers relations and ideas quickly			
5	Thinks of more than one solution to one problem			
6	Prefers to work with few instructions from the teacher			
7	Likes to construct things and situations in a new way			
8	Takes interest in evaluating things and events to promote them			
9	Produces a large number of ideas and solutions to problems			
10	Reaches unique and creative solutions to problems			
11	Is imaginative			
12	Criticizes constructively			
13	Proposes new techniques for classroom activities			
14	Asks interesting questions			
15	Possesses the ability to organize his ideas			
16	Is able to present new ideas			
17	Thinks deeply in topics and problems			
18	Pursues the work in hand with personal motivation			

19	Takes interest in bettering his work			
20	Is highly curious			
21	Takes risk and tries unfamiliar solutions			
22	Seeks superiority and excellence			
23	Completes the work assigned to him no matter how long it takes			
24	Prefers to read books that tackle religious, political, social…topics that appeal to the adults			
25	Takes interest in reading biographies, encyclopedias and atlases			
26	Prefers to work alone			
27	Grasps humor and jokes that other mates fail to grasp			
28	Gives humorous comments			
29	Admires aesthetic aspects in things			
30	Deals with ambiguous problems competently			
31	Adapts easily to new situations			
32	Expresses his opinion even if it contradicts others' opinion			
33	Has the ability to convince others with his opinion			
34	Carries responsibility well			
35	Can reconcile his mates' contradicting points of view			
36	Directs the group to make sound decisions about problems			
37	Takes the initiative in various situations			