Published in final edited form as:

J Psychoeduc Assess. 2008 September ; 26(3): 274–286. doi:10.1177/0734282907303873.

Measurement Invariance of the Chinese Gifted Rating Scales: Teacher and Parent Forms

Yaacov Petscher and Huijun Li

Florida State University Florida Center for Reading Research

Abstract

The Gifted Rating Scales–School Form (GRS-S) has been validated in several countries; however, no study has examined the rater invariance of this measure. The present study built on previous validity studies and examined configural and metric invariance between parent and teacher raters using the Chinese version of the GRS-S Teacher and Parent Forms. Analyses were conducted from a Jöreskog tradition, with testing of hypotheses related to group invariance. Model fit indices (i.e., comparative fit index [CFI], root mean square error of approximation [RMSEA], χ^2/df) supported equivalence in factor structure and standardized loadings between raters. Additionally, hypothesis testing revealed equivalence in loadings for raters. Implications of the results, as well as limitations and directions for future research, are discussed.

Keywords

Gifted Rating Scales; Chinese; rater invariance

China has had a long history of seeking out and educating child prodigies (Tsui-Chen, 1961). The main criteria in identifying a gifted child in China until the early 1900s had been literary creation (ability to compose poems and essays), reading, and exceptional memory of literary works. Recently, intelligence scores obtained from the Chinese versions of the Stanford-Binet Intelligence Test, the Wechsler Preschool and Primary Scale of Intelligence–Revised, and Test for Identifying the Cognitive Ability of Supernormal Children (TICASC) have been used to identify gifted children (Shi & Zha, 2000; Stevenson, 1998). Shi and Zha (2000) have suggested a broader definition of giftedness to encompass multiple dimensions, such as high cognitive ability, academic skills, leadership, creative potential, and even positive personality traits. Many of the aspects of giftedness are consistent with the recent definition of giftedness in other countries such as the United States. For example, the U.S. Department of Education has conceptualized giftedness as extraordinary intellectual and specific academic or artistic abilities or high performance capability in creativity or leadership (U.S. Department of Education, Office of Educational Research and Improvement, 1993).

To conduct a comprehensive evaluation of a potentially gifted child, teachers and parents are often invited to provide information on the child (Pfeiffer, 2002). Although early reports (e.g., Gear, 1978; Pagnato & Birch, 1959) indicated that teachers were relatively poor at identifying gifted students because they were too close to students to have a sound perspective, more researchers (e.g., Gagné, 1994; Hunsaker, Finley, & Frank, 1997; Renzulli & Delcourt, 1986), however, have argued that teacher nomination and teacher rating for gifted children identification were very useful in identifying gifted students. Furthermore, teacher accuracy can be improved through the use of psychometrically sound rating scales that have been formulated from a list of characteristics of gifted children. Therefore, well-designed teacher rating scales provide teachers with a reliable and efficient way to summarize their perception of a student based on a large number of classroom observations across a broad sample of academic tasks. Teacher rating scales are particularly valuable in the assessment of the various aspects of giftedness because there are very few direct assessment measures of creativity, artistic ability, and leadership (Jarosewich, Pfeiffer, & Morris, 2002).

Parent input has been another source of information in the identification of gifted students. Some researchers, however, indicate that parents may not be able to provide accurate ratings of their sons' or daughters' level of giftedness, compared to teacher ratings (Pfeiffer, 2002). The theory supporting this viewpoint is that many parents tend to inflate or exaggerate their child's abilities. Nevertheless, more researchers argue that parents can provide valuable and accurate information about their children that is not readily observable at school settings (e.g., Gogel, McCumsey, & Hewett, 1985; Gottfried, Gottfried, Bathurst, & Guerin, 1994). Parents are able to identify special talents in their children, especially when using research-supported instruments rather than locally designed checklists and observation forms (Gogel et al., 1985; Jin & Feldhusen, 2000).

China is known for its dedication and effort in screening, identifying, and educating its most talented students. Shi and Zha (2000) outlined principles for identifying the gifted students in China. The principles include a comparison of the IQ level of children of the same age, quantitative and qualitative analysis of responses to standard tests, an investigation of personality traits and physical fitness, and the actual performance of a child in an intellectually challenging classroom or school setting. In practice, typically, students' parents in China first submit an application to the testing centers, which notify the public of when a gifted assessment will take place. Once registered, the child takes a Chinese version of the Stanford-Binet Intelligence Test or the Wechsler Preschool or Primary Scale of Intelligence-Revised. Students who have obtained an IQ of 125 or higher will take the TICASC (Cooperative Research Group of Supernormal Children of China, 1985). Information on the child's personality traits and physical status are collected for students who obtain a score on the TICASC two or more standard deviations above the mean, or a score above the mean for children 2 years older, or a score greater than the 95th percentile for same-age children. A unique procedure in the assessment of giftedness in China is that students considered to be emotionally and behaviorally immature or having certain physical challenges may not be admitted into the gifted program even if they have passed the IQ and TICASC (Shi & Zha, 2000). The rationale is that these students may not be emotionally, behaviorally, or physically capable of coping with the intellectual and academic challenges

in the environment created for gifted children. Last, students who are successfully placed in a gifted class will be observed for several weeks in terms of their learning processes. If a student is able to take the intellectual challenges, he or she is permitted to remain in the gifted classroom (Shi & Zha, 2000).

Shi and Zha (2000) have indicated that there are several kinds of gifted education in China. First, there are about 50 primary and high schools across the country that offer gifted experimental classes. Second, those children who have passed the different exams are allowed to enter primary school, middle school, high school, or university earlier than the regular children, or to skip classes. Third, several universities offer “youth college class” for gifted children starting from age 12 or 13. These adolescents can get a college degree at age 16 or 17. Fourth, the gifted students sometimes receive education along with regular students, but in the meantime they are instructed individually and take advanced coursework in and out of school.

To assist the screening procedure by employing psychometrically sound instruments, we recently validated a gifted rating scale (Chinese Parent/Teacher Gifted Rating Scales–School Form) in China measuring intelligence, academic skills, leadership, artistic talent, creative ability, and level of motivation, which holds promise to help identify gifted students. The validation study of the Chinese teacher rating scales showed strong internal consistency and construct validity (Li, Pfeiffer, Petscher, Kumtepe, & Mo, in press). The original English version of the Gifted Rating Scales–School Form (GRS-S) was developed as a teacher rating scale, but it has been validated as a parent scale in different cultures where results showed strong internal consistency and construct validity (Lee & Pfeiffer, 2006; Li, Lee, Pfeiffer, & Petscher, 2007; Petscher & Pfeiffer, 2007).

In conducting the assessment of giftedness and other special needs of a child, a best practice encompasses collecting comprehensive information about the child from different informants such as teachers and parents, using fair, useful, and valid assessment instruments (Jacob & Hartshorne, 2007). Comparability of measures is essential for valid research across rating sources and may be attained by demonstrating the equivalence of constructs across raters. Social science researchers are increasingly concerned with testing for measurement invariance because of its critical importance when comparing between-groups/interrater differences. Some researchers have argued that

if measurement invariance cannot be established, the between-group or interrater differences cannot be meaningfully interpreted because one will not know whether this difference is because of a “true” difference on the construct of interest or to different psychometric responses to the items. (Ang, Huan, & Braman, 2007, p. 76)

The scope of our previous validation studies of the Gifted Rating Scales (GRS-S) did not allow further analyses of measurement invariance of teacher and parent ratings. The purpose of this study was to extend previous validation research and examine factor invariance across raters to inform the field if teachers and parents share similar perceptions of giftedness in children.

METHOD

Participants and Procedure

The study was conducted in two elementary schools and one middle school in Southwest China. Fifty teachers participated in this study and provided ratings on 500 students. One rating form was dismissed because of missing data, yielding a total of 499 student-rating forms. In China, students have more than one teacher; however, there is one primary teacher who oversees the major academic content areas and is most familiar with each student. These primary teachers were invited to select 10 students in their respective classes (average class size was 40 students), with 2 students functioning much above average level, 2 students functioning above average level, 2 students functioning at the average level, 2 students functioning below average level, and 2 students functioning much below average level.

Of the 499 students, 224 parents participated in this study and rated their children independently. This yielded 224 parent rating forms, all from mothers of the children. The average age of mothers was 38.12 ($SD = 4.09$). The mothers' level of education ranged from completion of high school to postgraduate education, with 44% from high school, 40% from college, 10% from postgraduate schools, and the rest not reporting their level of education. Employment information revealed that the mothers' jobs fell into the following categories: A total of 23.2% of the mothers held government positions, 34.3% were educational professionals, 12.5% were general laborers, and about 30% were employed as business people, technicians, medical personnel, and clerks. Among the 224 students, 126 students were boys (56.3%) and 98 were girls (43.8%), and their ages ranged from 7 to 13 years ($M = 10.02$, $SD = 1.95$). The grade distribution for the sample was as follows: first grade (13.4%), second grade (20.5%), third grade (11.6%), fourth grade (12.5%), fifth grade (20.1%), sixth grade (13.4%), and seventh grade (8.5%).

Instrumentation

The Chinese GRS–Teacher and Parent Forms—The Chinese GRS–Teacher and Parent Forms were directly adapted from the GRS–S Form (Pfeiffer & Jarosewich, 2003). Both forms are identical, except that the word *student* has been changed to *child* in the directions section of the parent form. This rating scale was translated into Chinese by the second author following a thorough procedure (Li et al., in press). The Chinese GRS–Teacher and Parent Forms are for children aged 6:0 to 13:11, and they consist of six scales (Intellectual Ability, Academic Ability, Creativity, Artistic Talent, Leadership Ability, and Motivation), with 12 items each for a total of 72 items. Below is a brief description of each of the six GRS–S scales:

Intellectual Ability—This scale measures the teacher's perception of a child's verbal and/or nonverbal mental skills, capabilities, or intellectual competence. Aspects of intelligence measured by this scale include abstract reasoning, problem solving, mental speed, and memory.

Academic Ability—This scale measures the teacher’s perception of a child’s skill in dealing with factual and/or school-related material. Advanced competence and high levels of proficiency in reading, math, and other aspects of the school curriculum are indicative of academic talent, as is facility in acquiring new knowledge and skills and the ability to understand complex material. Students who are academically gifted often have large stores of information including broad knowledge of the world around them.

Creativity—This scale measures the teacher’s perception of a child’s ability to think, act, and/or produce unique, original, novel, or innovative thoughts or products. Creativity can be expressed in a variety of ways: how a student solves problems, experiments with new ideas, formulates a solution to a group project, and/or uses imagination. Creatively gifted children are inventive, curious, and inquisitive. They show a preference for challenge and complexity and engage in problem solving that is insightful and creative.

Artistic Talent—This scale measures the teacher’s perception of a child’s potential for, or evidence of ability in, drama, music, dance, drawing, painting, sculpture, singing, playing a musical instrument, and/or acting. Artistic talent can be expressed in a variety of ways: how a child approaches activities, completes assignments, and/or uses art supplies or artistic media. Artistically gifted children learn artistic skills quickly and exhibit more technical sophistication.

Leadership Ability—This scale measures the teacher’s perception of a child’s ability to motivate others toward a common or shared goal. Leaders understand and monitor social dynamics and have strong interpersonal communication and conflict-resolving skills. They effectively orchestrate collective action and influence group behavior. Gifted leaders exhibit strong social judgment and prosocial values such as integrity and trustworthiness and demonstrate the drive to lead and take initiative.

Motivation—This scale refers to the child’s drive or persistence, desire to succeed, tendency to enjoy challenging tasks, and ability to work well without encouragement or reinforcement. The Motivation scale is not viewed as a type of giftedness but rather as the dynamic energy that drives or impels a student to achieve, akin to Tannenbaum’s (2000) formulation that emphasizes drive as an important factor in explaining a gifted child’s performance.

Each item is rated on a 9-point scale divided into three ranges: 1 to 3 (*below average*), 4 to 6 (*average*), and 7 to 9 (*above average*). The rating scales allow teachers and parents to determine first whether the child is below average, average, or above average for each item, compared to other students of the same age. The rating scales then allow the teachers and parents to determine, within the specific range, whether the child is at the bottom, middle, or top of the range. The standard (not translated/English) version of the GRS-S reports high levels of internal consistency (.97-.98), small structural equation modeling (SEM) and confidence intervals (CIs), and strong evidence of validity (Pfeiffer & Jarosewich, 2003; Ward, 2005). The coefficient alphas of the Chinese-translated GRS-S parent ratings ranged from .94 to .96 (Li et al., 2007), and those of the teacher ratings were between .95 and .99

(Li et al., in press). Further information about the GRS-S scales can be found in Pfeiffer and Jarosewich (2003).

Statistical Analyses

Building on previous work by Li et al. (in press) and Li et al. (2007), the measurement models of both teacher and parent responses were tested to estimate the invariance of the factor structure across groups. To this end, three separate groups of measurement models were tested to (a) separately estimate a baseline model for the teacher and parent groups to establish the equivalency of the theoretical factor structure (confirmatory factor analysis [CFA]), (b) test the invariance of factor loadings (multiple-group SEM), and (c) test the invariance of the factor covariances (multiple-group SEM; Byrne, 1998).

CFA is a theory-testing, direct-effects model that allows the researcher to form hypotheses about the relationships of observed indicators to latent constructs. Latent variables are typically described a priori, as opposed to exploratory factor analysis, which seeks to mathematically determine the most suitable number of components or factors that account for between-factors variability. Additionally, a second-order CFA is an appropriate methodology to test for the presence of a global factor(s), when the correlations among the latent constructs are, themselves, of a moderate or strong nature. This concept can be applied to as many levels as necessary (third order, fourth order, etc.). Many authors studying aptitude have proposed that intelligence can be viewed as unidimensional (Jensen, 1999), whereas others contended that intelligence is multidimensional (Gardner, 2000). Because the GRSS is based on a multidimensional model of giftedness, it was hypothesized that the scales could be better explained by second-order constructs.

All measurement models were tested using LISREL 8.1 (Jöreskog & Sörbom, 1993) and analyzing covariance matrices using a series of estimation methods (i.e., maximum likelihood, two-stage least squares) appropriate for the three identified objectives. Models were evaluated using a series of selected indices provided by LISREL, including the comparative fit index (CFI; Hu & Bentler, 1998), normed fit index (NFI), goodness-of-fit index (GFI), root mean square error of approximation (RMSEA), and the ratio between the model chi-square and degrees of freedom (χ^2/df). Values greater than .90 are typically indicative of good fit for the CFI, NFI, and GFI, whereas RMSEA values less than .05 signify good fit and are considered to be acceptable up to .10 (Hu & Bentler, 1999). Because the chi-square is sensitive to changes in the covariance matrix when samples are greater than 200, most researchers opt to use the χ^2/df ratio as an index of model fit, with results less than 3.0 indicating good model fit (Marsh, Balla, & McDonald, 1988).

Validity of a Baseline Model

CFA analyses for the validity of the teacher baseline model were derived from our previous results that indicated a six-factor model (Teacher–first order, six factors; Table 1) fit the data, compared to a unidimensional model (Teacher–first order, one factor; Table 1; Li et al., in press). Given the high correlations across factors that were consistently observed across GRS-S validation studies (Lee & Pfeiffer, 2006; Pfeiffer & Jarosewich, 2003), it was of interest to test a series of hierarchical factor models. These included using the six first-

order factors as indicators of one, two, or three second-order constructs, to partial out the collinearity among scales.

It was hypothesized that the Intellectual Ability and Academic Ability scales (which appeared to be scholastically oriented), the Creative and Artistic Talent (innovatively oriented), and Leadership Ability and Motivation (behaviorally oriented) pairs could be better explained by higher order factors (three second-order factors), because the theoretical relationships within each set are similar and conceptually measure aspects of a similar component. Furthermore, a unidimensional second-order factor model was hypothesized to fit as well as the first-order, six-factor model. In theory, this unidimensional component would represent a general intelligence (i.e., *g*) factor, whereby the individual scales would be indicating a rating of overall aptitude. Last, a two-factor hierarchical model was also proposed by grouping the behaviorally oriented scales together (i.e., Leadership Ability and Motivation) and the remaining four scales as a group to test the relative factor differences in behavior and nonbehavior domains.

Testing for Configural and Metric Invariance

The configural invariance approach is based on Thurstone's principle of simple structure (Thurstone & Thurstone, 1941). In essence, this principle states that the pattern of nonzero and near-zero loadings defines the structure of the measurement. Configural invariance requires that the same variables load on the same number and pattern of factors across groups (de Frias & Dixon, 2005). Metric invariance examined the degree of factor-loading equivalence across parent and teacher ratings. Methodology for the invariance testing involved a process of estimating a measurement model for each group independently, testing for configural invariance across groups, fitting the multiple-group model under the null and alternative hypotheses (i.e., specification of model), performing a chi-square difference test between the null and alternative models, testing the metric invariance, and performing a chi-square difference test for the null and alternative metric models (Du Toit & Du Toit, 2001).

RESULTS

Descriptive statistics for parent and teacher response are reported in Table 2. As expected, the data were normally distributed with a mean approximately equal to 50, and standard deviations centered approximately on 10. These values were comparable to findings observed in the original U.S. standardization sample (Pfeiffer & Jarosewich, 2003). Furthermore, little skew (range = $-.148$ to $.726$) or kurtosis (range = $-.916$ to $-.216$) was observed. Correlations among the scales for the original six-factor model are also reported in Table 2, with teacher estimates replicating previous findings in the standardization sample and parent estimates approaching values found in the English version of the parent form (Petscher & Pfeiffer, 2007).

Individual Model Testing

Teacher models—As mentioned earlier, prior testing of the factor structure of the Chinese GRS–Teacher Form (Li et al., in press) indicated that a six-factor solution (i.e., Teacher–first order, six factors) provided better fit to the data than a unidimensional

measurement model. Findings from the teacher *g* factor model (i.e., Teacher–second order, one factor; Table 1) indicated that a high degree of error was observed (RMSEA = .133, root mean residual [RMR] = .07) in the model. Although the relative fit indices (i.e., RFIs, NFI) suggested the model fit well to the data, the overall fit of the model was not acceptable (e.g., large error, poor absolute fit). Results from the Teacher second-order model with two second-order factors indicated that the model was an improvement over the prior model, $\chi^2(2, 475) = 9,461.20, p < .001, \chi^2/df = 3.82, RMSEA = .088$ (90% CI = .086, .089), NFI = .98, CFI = .98, RFI = .97, and RMR = .13. Although the loadings in this model were significantly different from zero, the relative fit for both the *g* factor and current model were not as good as the original six-factor model was to the data. Results from the Teacher second-order, three-factor model indicated that this model provided fit information comparable to the Teacher first-order, six-factor solution; the fit was stronger (χ^2/df , RMSEA) or similar (NFI, CFI, RFI, and RMR) in all areas for the latter model, though the differences between the two were negligible.

Parent models—Similar to the findings from the initial teacher form testing, the proposed six-factor solution for the GRS–Parent Form (Li et al., 2007) also exhibited good model fit: $\chi^2(2, 469) = 5,537.63, p < .001, \chi^2/df = 2.24, RMSEA = .076$ (90% CI = .074, .079), NFI = .96, CFI = .98, RFI = .98, and RMR = .06. The addition of three second-order factors, similar to the permutations of the Teacher Form was considered; however, given the results from the teacher analyses, the only hierarchical model tested for the parents was a three second-order model. Results from this analysis approximated that of the teacher analyses, that is, the nested second-order-factor model fit as well as the original six-factor solution $\chi^2(2, 475) = 5,575.73, p < .001, \chi^2/df = 2.25, RMSEA = .078$ (90% CI = .075, .080), NFI = .96, CFI = .98, RFI = .98, and RMR = .06. Differences in the factor correlation matrix between the two measurement solutions were negligible, with the largest observed correlation remaining between Intellectual Ability and Academic Ability ($r = .86$).

Although second-order factors would be important for estimating latent factor scores, the goal of the present study was focused more on measurement structure analysis that would be useful to researchers and practitioners alike. Because both the six-factor and the three-factor models fit the data well, the authors decided to use the six-factor solution to serve as the base model for configural and invariance analyses, given the prior theoretical work (Lee & Pfeiffer, 2006; Pfeiffer & Jarosewich, 2003). A summary of fit indices for the parent and teacher structural models is reported in Table 1.

Rater Invariance Testing

Configural invariance for the six-factor model between teacher ($n = 499$) and parent ($n = 224$) ratings fit the data reasonably well (Table 3), $\chi^2(5, 097) = 17,263.07, p < .001, \chi^2/df = 3.38, RMSEA = .087$ (90% CI = .086, .088), CFI = .98, with all factor loadings larger than .40 being significantly larger than 0 ($z > 1.96$). Hypothesis testing of the measurement model (Table 4) indicated that the six-factor solution was not supported by data from the different samples. Metric invariance of factor loadings did not significantly differ in fit from the configural model, $\chi^2(5, 092) = 17,248.59, p < .001, \chi^2/df = 3.38, RMSEA = .087$ (90% CI = .086, .088), CFI = .98, $\chi^2 = 14.48, df = 5, p = .48$. Testing of the null hypothesis for the

metric invariance of loadings (Table 4) indicated that evidence existed for the equivalence of factor loadings across raters.

DISCUSSION

The purpose of this study was to further evaluate the factor structure of the Chinese GRS–Teacher and Parent Forms and examine the invariance of parameter estimates between parent and teacher ratings of student abilities.

Individual Model Testing

Reevaluation of the Chinese GRS–Teacher Form indicated that a six-factor solution was best fitting the data, although iterative testing of second-order factor models revealed that a six-factor model with three hierarchical factors fit nearly as well as the original measurement model. Although the six-factor model was retained for invariance testing, it does not rule out the fact that a hierarchical model may represent a more parsimonious description of construct relationships. Furthermore, a hierarchical model could potentially be more useful in the construction of practical factors scores when screening students for gifted abilities.

Sequential testing of the parent measurement model indicated that the six-factor solution was appropriately identified. A review of the item–scale relationships showed that moderate to very strong correlations were observed, with a large proportion of the variance in the item explained well by the proposed factor. Although the fit indices suggested that the model chi-square could be reduced by cross-loading three items from the Intellectual Ability factor on the Academic Ability factor, two items from the Academic Ability factor to Intellectual Ability, and two items from Artistic Talent to Creativity, further investigation of these relationships is warranted before the removal of any item. Though recent validation studies have shown strong correlations among the proposed six-factor solution (Lee & Pfeiffer, 2006; Li et al., in press; Pfeiffer & Jarosewich, 2003; Rosado, Pfeiffer, & Petscher, 2007), stronger statistical models should be employed and judged along with theoretical concerns before item removal. Generally, the model fit results provide information that scores from the Chinese GRS–Parent Form retain strong psychometric properties and may be useful for teachers and psychologists in the screening of gifted Chinese students.

Configural and Metric Invariance

The measurement invariance of the six-factor solution across raters was supported with strong fit estimates ($CFI = .98$, $RMSEA = .087$) and little change between the configural and metric model tests. Measurement equivalence between raters provided evidence that the meaning of giftedness defined by the scales of the Chinese GRS–Teacher and Parent Forms is comparable across parent and teacher raters. One reason for this may be that most Chinese parents at the current time have only one child in the family and they are heavily, if not overly, involved in their children’s schoolwork. Parents regularly communicate with the teachers regarding their child’s academic and other school-related progress, and teachers also discuss with parents the different talents that their child has and the challenges they face and what parents need to do to develop the talents and overcome the challenges. Thus, it

may not be surprising that teachers and parents share similar perceptions of giftedness in children (Xiaoti Li, personal communication, March 31, 2007).

General Conclusion

Prior validation work on Chinese GRS–Teacher and Parent Forms (Li et al., 2007, in press) provided information relative to the internal consistency, discriminant validity of factors, and predictive validity of the subscales for a large target sample. The current study established a fuller understanding of particular measurement components of the model by examining the underlying structure of gifted abilities as rated by both teachers and parents. Because these two groups represent those who have the most contact with children, it is important to establish that the perception of giftedness across raters is invariant to provide a comprehensive and reliable screening method for practitioners.

Results from the second-order factor models were particularly interesting in the context of the theory that describes general intelligence compared to multiple intelligences. Findings from the current study support a multidimensional intelligence model, evidenced by the acceptable model fit in both the first-order six-factor solutions and three second-order six-factor solutions. Although the *g* model displayed a trend toward model fit, one cannot ignore the high degree of error and poor absolute fit compared to the other two models. One might argue that Motivation and Leadership Ability are not forms of intelligence, thereby convoluting the *g* factor model; however, the two-factor hierarchical model did parse out the behaviorally oriented constructs, and this did not fit as well as the further developed three-factor hierarchical model.

Future studies should examine the extent to which ratings are reliable over time, as well as the degree of convergence of ratings across factors (i.e., multitrait–multimethod design). Furthermore, given the degree of the potential reduction in the model chi-square for a cross-loading of items, more sophisticated statistical models could be employed to evaluate the information functions of the items. The utility of Samejima’s (1972) graded response model could result in a shorter, more concise version of the Chinese GRS–Teacher and Parent Forms.

A few limitations of the study should be brought to the attention of the reader. The high correlations among the scales of both the parent and teacher forms may be from a resulting halo effect, that is, it was conceivable that teachers’ and parents’ view of a child’s ability (*g*) influenced how they viewed other domains of rating. As an example, the relationship between Motivation and Intellectual Ability on the Teacher Form was .84, and .89 with Academic Ability. These correlations are quite high, given little empirical support for high correlations between cognitive abilities and motivation (Gagné & St. Père, 2001). Conversely, the high correlations between Intellectual Ability and Academic Ability on both the Teacher (.93) and Parent (.83) Forms were consistent with prior research between ability and achievement measures in the general population (Sattler, 2001; Wechsler, 2003). Additionally, sampling participants from teacher-rated categories of academic achievement may have moderated both the magnitudes and variability of the ratings. Future research should use more representative sampling techniques and stratify students independent of teacher nominations.

References

- Ang RP, Huan VS, Braman OR. Factorial structure and invariance of the academic expectations stress inventory across Hispanic and Chinese adolescent samples. *Child Psychiatry and Human Development*. 2007; 38:73–87. [PubMed: 17131177]
- Byrne, B. Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming. Lawrence Erlbaum; Mahwah, NJ: 1998.
- Cooperative Research Group of Supernormal Children of China. Test for identifying the cognitive ability of supernormal children. Author; Beijing: 1985.
- de Frias CM, Dixon RA. Confirmatory factor structure and measurement invariance of the Memory Compensation Questionnaire. *Psychological Assessment*. 2005; 17:168–178. [PubMed: 16029104]
- Du Toit, M.; Du Toit, S. Interactive LISREL: User's guide. Scientific Software International; Lincolnwood, IL: 2001.
- Gagné F. Are teachers really poor talent detectors? Comments on Pagnato and Birch's (1959) study of the effectiveness and efficiency of various identification techniques. *Gifted Child Quarterly*. 1994; 38:124–126.
- Gagné F, St. Père F. When IQ is controlled, does motivation still predict achievement? *Intelligence*. 2001; 30:71–100.
- Gardner, H. Intelligence reframed: Multiple intelligence for the 21st century. Basic Books; New York: 2000.
- Gear G. Effects of training on teachers' accuracy in the identification of gifted children. *Gifted Child Quarterly*. 1978; 22:90–97.
- Gogel EM, McCumsey J, Hewett G. What parents are saying. *G/C/T*. 1985; 41:7–9.
- Gottfried, AW.; Gottfried, AE.; Bathurst, K.; Guerin, DW. The Fullerton longitudinal study. Plenum; New York: 1994. Gifted IQ: Early developmental aspects.
- Hu L, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*. 1998; 3:424–453.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.
- Hunsaker SL, Finley VS, Frank EL. An analysis of teacher nomination and student performance in gifted programs. *Gifted Child Quarterly*. 1997; 41:19–24.
- Jacob, S.; Hartshorne, TS. Ethics and law for school psychologists. John Wiley; Hoboken, NJ: 2007.
- Jarosewich T, Pfeiffer SI, Morris J. Identifying gifted students using teacher rating scales: A review of existing instruments. *Journal of Psychoeducational Assessment*. 2002; 20:322–336.
- Jensen AR. The G factor: The science of mental ability. *Psychology*. 1999; 10 #23 Intelligence G Factor (1). from <http://psycprints.ecs.soton.ac.uk/archive/00000658/>.
- Jin SU, Feldhusen JF. Parent identification of the talents of gifted students. *Gifted Education International*. 2000; 23:230–236.
- Jöreskog, KG.; Sörbom, D. LISREL 8: User's reference guide. Scientific Software International; Chicago: 1993.
- Lee D, Pfeiffer SI. The reliability and validity of a Korean translated version of the Gifted Rating Scale. *Journal of Psychoeducational Assessment*. 2006; 24:210–224.
- Li, H.; Lee, D.; Pfeiffer, SI.; Petscher, Y. Parent rating using Gifted Rating Scale-School Form: Reliability and validity in Chinese students. 2007. Manuscript submitted for publication
- Li H, Pfeiffer SI, Petscher Y, Kumtepe AK, Mo G. Validation of the Gifted Rating Scales-School Form in China. *Gifted Child Quarterly*. (in press).
- Marsh HW, Balla JR, McDonald RP. Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*. 1988; 103:391–410.
- Pagnato CW, Birch JW. Locating gifted children in junior high schools: Comparison of methods. *Exceptional Children*. 1959; 25:300–304.
- Petscher, Y.; Pfeiffer, SI. Preliminary examination of the reliability and validity of the Gifted Rating Scales-Parent Form. 2007. Manuscript submitted for publication

- Pfeiffer SI. Identifying gifted and talented students: Recurring issues and promising solutions. *Journal of Applied School Psychology*. 2002; 1:30–50.
- Pfeiffer, SI.; Jarosewich, T. *Gifted Rating Scales*. PsychCorp, Harcourt Assessment; San Antonio, TX: 2003.
- Renzulli JS, Delcourt MAB. The legacy and logic of research on the identification of gifted persons. *Gifted Child Quarterly*. 1986; 30:20–23.
- Rosado, J.; Pfeiffer, SI.; Petscher, Y. The reliability and validity of a Spanish translated version of the Gifted Rating Scales. 2007. Manuscript submitted for publication
- Samejima F. A general model for free response data. *Psychometrika*. 1972; 37:1–18.
- Sattler, JM. *Assessment of children: Cognitive applications*. 4th ed. Author; San Diego, CA: 2001.
- Shi, J.; Zha, Z. Psychological research on and education of gifted and talented children in China. In: Heller, KA.; Monks, FJ.; Sternberg, RJ.; Subotnik, RF., editors. *International handbook of giftedness and talent*. 2nd ed. Elsevier; Oxford, UK: 2000. p. 757-764.
- Stevenson, HW. Cultural interpretations of giftedness: The case of East Asia. In: Friedman, RC.; Rogers, KB., editors. *Talent in context*. American Psychological Association; Washington, DC: 1998. p. 61-77.
- Tannenbaum, AJ. A history of giftedness in school and society. In: Heller, KA.; Monks, FJ.; Sternberg, RJ.; Subotnik, RF., editors. *International handbook of giftedness and talent*. 2nd ed. Elsevier; Oxford, UK: 2000. p. 23-51.
- Thurstone, LL.; Thurstone, T. *Factorial studies of intelligence (Psychometric Monograph 2)*. University of Chicago Press; Chicago: 1941.
- Tsui-Chen, O. Some facts and ideas about talent and genius in Chinese history. In: Bureday, GZF.; Lauwerys, A., editors. *Concepts of excellence in education: The yearbook of education*. Harcourt, Brace and World; New York: 1961. p. 54-61.
- U.S. Department of Education, Office of Educational Research and Improvement. *National excellence: A case of developing America's talent*. Author; Washington, DC: 1993.
- Ward, SA. Gifted Rating Scales. In: Spiers, RA.; Plake, BS., editors. *16th mental measurements yearbook*. University of Nebraska Press; Lincoln: 2005. p. 404-407.
- Wechsler, D. *Wechsler Intelligence Scale for Children*. Fourth edition. Psychological Corporation; San Antonio, TX: 2003.

Table 1

Fit Indices of Individual Measurement Models

Model	χ^2	df	RMSEA (90% CI)	NFI	CFI	RFI	RMR
Teacher - first order							
One factor	20,272.96	2,484	.210 (.200, .220)	.97	.97	.97	.06
Six factor	7,269.84	2,469	.077 (.075, .078)	.99	.99	.99	.03
Teacher - second order							
One second-order factor	13,630.22	2,478	.133 (.131, .135)	.98	.97	.98	.07
Two second-order factors	12,169.53	2,478	.130 (.121, .135)	.98	.98	.97	.06
Three second-order factors	8,349.73	2,475	.082 (.080, .084)	.99	.99	.99	.03
Parent - first order							
Six factor	5,537.63	2,469	.076 (.074, .079)	.96	.98	.98	.06
Parent - second order							
Three second-order factors	5,575.73	2,475	.078 (.075, .080)	.96	.98	.98	.06

Table 2

Correlation, Descriptive, and Internal Consistency Estimates for GRS-Chinese

	1	2	3	4	5	6	7	8	9	10	11	12
1. Intellectual - T ^a												
2. Academic - T	.93											
3. Creativity - T	.90	.91										
4. Artistic - T	.78	.83	.82									
5. Leadership - T	.79	.85	.78	.79								
6. Motivation - T	.84	.89	.82	.79	.91							
7. Intellectual - P ^b												
8. Academic - P							.83					
9. Creativity - P							.76	.71				
10. Artistic - P							.64	.68	.75			
11: Leadership - P							.73	.75	.74	.70		
12. Motivation - P							.71	.75	.63	.61	.74	
Mean	50.06	49.61	50.76	51.63	47.86	49.86	50.44	51.70	52.59	52.59	51.31	52.57
Standard deviation	13.35	12.67	12.73	13.78	12.14	12.26	8.84	9.56	11.35	11.35	9.62	9.88
Internal consistency	.97	.98	.95	.99	.97	.98	.94	.96	.96	.96	.92	.96

Note: GRS = Gifted Rating Scales.

^aTeacher form.

^bParent form.

Table 3

Invariance of Chinese GRS–Teacher and Parent Forms

Model	χ^2	df	RMSEA (90% CI)	CFI	χ^2	df
Configural	17,263.07	5,097	.087 (.086, .088)	.98		
Metric	17,258.59	5,092	.087 (.086, .088)	.98		
					4.48	5

Note: GRS = Gifted Rating Scales; RMSEA = root mean square error of approximation; CI = confidence interval; CFI = comparative fit index.

Table 4

Rater Hypothesis Testing—Teacher and Parent Forms

Hypothesis	χ^2	df	p Value
Measurement invariance			
H0	19,213.59	5,097	
H1	16,597.71	4,959	
Difference	2,615.876	138	.00
Loading invariance			
H0	16,691.44	5,032	
H1	16,597.71	4,959	
Difference	93.729	73	.05