Psicothema

Psicothema 2014, Vol. 26, No. 3, 423-430
doi: 10.7334/psicothema2013.186

ISSN 0214 - 9915 CODEN PSOTEG

www.psicothema.com

Psychometric properties of the Spanish Adaptation of the School Attitude Assessment Survey-Revised

Pablo Miñano Pérez, Juan Luis Castejón Costa and Raquel Gilar Corbí
Universidad de Alicante

Abstract

Background: The assessment of attitudes toward school with the objective of identifying adolescents who may be at risk of underachievement has become an important area of research in educational psychology, although few specific tools for their evaluation have been designed to date. One of the instruments available is the School Attitude Assessment Survey-Revised (SAAS-R). **Method:** The objective of the current research is to test the construct validity and to analyze the psychometric properties of the Spanish version of the SAAS-R. Data were collected from 1,398 students attending different high schools. Students completed the SAAS-R along with measures of the g factor, and academic achievement was obtained from school records. **Results:** Confirmatory factor analysis, multivariate analysis of variance and analysis of variance tests supported the validity evidence. **Conclusions:** The results indicate that the Spanish version of the SAAS-R is a useful measure that contributes to identification of underachieving students. Lastly, the results obtained and their implications for education are discussed.

Keywords: Instrument cross-validation, high school student, underachievement, confirmatory factor analysis, reliability.

Resumen

Propiedades psicométricas de la adaptación española de la escala School Attitude Assessment Survey-Revised. **Antecedentes:** la evaluación de las actitudes hacia la escuela con el objetivo de identificar adolescentes que pueden estar en riesgo de tener un rendimiento menor a lo esperado en un área importante de investigación en psicología de la educación, aunque se han diseñado pocos instrumentos de evaluación hasta la fecha. Uno de los instrumentos disponibles es el School Attitude Assessment Survey-Revised (SAAS-R). **Método:** el objetivo de este trabajo es examinar la validez de constructo y analizar las propiedades psicométricas de la versión española de la Escala Revisada de Evaluación de Actitudes hacia la Escuela (SAAS-R). Los datos se obtuvieron de 1.398 estudiantes que asistían a diferentes institutos de Educación Secundaria. Los estudiantes cumplimentaron el SAAS-R junto con medidas del factor g, y se obtuvo el rendimiento académico de las actas escolares. **Resultados:** los análisis factorial confirmatorios, el análisis multivariado de la varianza y las pruebas de análisis de varianza sustentaron la evidencia de validez. **Conclusiones:** los resultados indican que la versión española del SAAS-R es una medida útil que contribuye a la identificación de estudiantes con rendimiento menor a lo esperado. Finalmente, se discuten los resultados y sus implicaciones educativas.

Palabras clave: validación cruzada, estudiantes de Educación Secundaria, rendimiento inferior al esperado, análisis factorial confirmatorio, fiabilidad.

The School Attitude Assessment Survey-Revised (SAAS-R) is a relatively new instrument that was initially developed by McCoach (2002) at the University of Connecticut, USA, with the objective of identifying adolescents who may be at risk of underachievement. The most common definition of underachievement characterizes it as a discrepancy between potential (or ability) and performance (or achievement) (Dowdall & Colangelo, 1982; McCoach & Siegle, 2003b).

The initial version of the School Attitude Assessment Survey (SAAS) measured four factors associated with underachievement: (a) Academic Self-Perception, which explored students' perception of their academic abilities; (b) Attitude toward School, which consisted of the students' self-reported interest in and affect toward school (McCoach & Siegle, 2003b); (c) Motivation, "students also must be motivated to use the strategies as well as regulate their cognition and effort" (Pintrich & DeGroot, 1990, p. 33); and (d) Peer Attitudes toward School, students' perceptions of how their peers value achievement. The SAAS instrument consisted of 20 questions and employed a 7-point Likert-type scale. Initial research with 1738 secondary students supported the correlation between the SAAS and academic achievement, measured as self-reported Grade Point Average (GPA). When the SAAS was administered in a cross-validation sample (n = 420 ninth-grade students from a multiethnic high school), the model fit was acceptable: $\chi^2_{(162)}$ = 509.5, p<.001, CFI = .94, TLI = .92, RMSEA = .075, SRMR = .045 (McCoach, 2002). The scores showed internal consistency reliability coefficients of at least .80 on each of the factors.

Received: June 5, 2013 • Accepted: April 23, 2014
Corresponding author: Pablo Miñano Pérez
Facultad de Educación
Universidad de Alicante
03080 Alicante (Spain)
e-mail: pablo.m@ua.es

Furthermore, preliminary results suggested that the SAAS could separate approximately 90% of high GPA and low GPA students into the correct classifications (McCoach, 2002). However, the results of this study also indicated a certain degree of overlap between some of the scales.

In order to improve the SAAS instrument, a revised version was developed, the SAAS-R (McCoach & Siegle, 2003b), in which the factor Peer Attitudes toward School was removed and two new factors were added: one to measure students' valuation of the school's goals, Goal Valuation, since children's goals and achievement values are key mediators of their academic self-regulation, and another, Attitudes toward Teachers, to separate this from students' general attitudes toward school, since the authors assumed that students may present positive affect toward the school for various reasons without necessarily having positive attitudes about their classes and teachers.

To carry out construct validation of the SAAS-R, a pilot version was developed with 48 questions to measure the five factors, which was administered to a sample of 942 high school students with diverse demographic characteristics and achievement levels. Following a succession of confirmatory factor analyses, the pilot version was eventually reduced to 35 questions in the final version of the instrument. Of these, 24 were questions from the original pilot version, 2 were reworded questions and 9 were newly written questions. Thus, the final instrument consisted of 7 questions on the Academic Self-Perceptions factor, 7 questions on the Attitudes toward Teachers factor, 5 questions on the Attitudes toward School factor, 6 questions on the Goal Valuation factor and 10 questions on the Motivation/Self-Regulation factor.

The final model supported the five-factor structure of the SAAS-R and exhibited a reasonable fit: $\chi^2_{(550)} = 1581.7$, CFI = .911, TLI = .918, RMSEA = .059, SRMR = .057 (n = 645). The results also evidenced a consistent intercorrelation (.74) between the Goal Valuation and Motivation/Self-Regulation factors. The SAAS-R showed acceptable reliability with an internal consistency for each scale above .85. In addition, several studies have analyzed criterion-related validity, confirming the correlation between attitudes measured by the SAAS-R and students' academic achievement (McCoach & Siegle, 2001, 2003a).

Subsequently, Suldo, Shaffer and Shaunessy (2008) conducted research on the psychometric properties of the SAAS-R, examining whether use of this scale could be extended beyond gifted students alone to students in general education and college preparatory programs. Participants were 321 gifted and non-gifted students aged between 14 and 19 years old (M = 15.72, SD = 1.24) attending a rural high school in a southeastern state in the USA. The sample included multiple ethnic groups. Confirmatory factor analysis, using item parcels rather than individual items, was conducted to validate the five-factor structure of the SAAS-R, obtaining a good fit according to the criteria: SRMR = .04, CFI = .96, TLI = .94 and RMSEA = .08. High levels of internal consistency reliability were also obtained for each of the factors, with alpha values ranging between .88 and .93.

Intercorrelations among SAAS-R scales were moderate to high, the highest being Attitudes toward School with Attitudes toward Teachers (.72) and Goal Valuation with Motivation/Self-Regulation (.63). Concurrent criterion-related validity was also examined and it was found that students with low achievement obtained significantly lower scores on each SAAS-R scale compared to students with high and average achievement. Convergent validity was assessed and confirmed by correlational analysis between individual SAAS-R scales and measures of other indicators of academic performance (Suldo et al., 2008).

The SAAS-R instrument has also been used in other studies to assess the gap between aspirations and expectations (Kirk et al., 2012), in research and development of interventions for gifted students (Rubenstein, Siegle, Reis, McCoach, & Burton, 2012), and at-risk adolescents (Lang, Waterman, & Baker, 2009) among others.

Given the above, the primary purpose of the current study was to determine the Spanish adaptation and psychometric properties of the SAAS-R in a Spanish sample. In addition, because most of the research on the SAAS-R has been conducted with samples of gifted students (Suldo et al., 2008), a further aim was to study a large and more heterogeneous sample.

The first objective was to develop a Spanish version of the SAAS-R, and examine the construct validity and reliability of this instrument in a representative sample of secondary students.

Since several studies have shown moderate to high overlap between some of the scales, with correlation coefficients of up to .72, along with a high consistency among items, the second objective was to conduct a comparison between a model with five first-order correlated factors and a model with five first-order factors and one second-order factor.

Once the instrument had been designed to measure factors distinguishing underachievers from normal achievers in a secondary school setting (McCoach & Siegle, 2003b), the third objective was to examine test-criterion relationships to test whether the scores obtained in the SAAS-R subscales would serve to differentiate between groups of academic underachievers, normal achievers and overachievers.

Method

Participants

A total of 1,456 students in their first and second year of compulsory secondary education in the province of Alicante (Spain) participated in this study. Of these, 58 were excluded from the final sample due to having an insufficient command of the language, not having completed the tests in their entirety or because they did not have parental consent. Thus, the final sample consisted of 1,398 subjects (n = 1,398).

Of these, 732 were enrolled in their first year (52.4%), while the remaining 666 were in their second year (47.6%). Of the sample, 52.8% were males and 47.2% females, and their ages ranged between 11 and 15 years old (M = 12.5, SD = 0.67).

Random cluster sampling was used to select the sample, using the school as the sampling unit. A total of eight high schools in the area participated, two of which were state-assisted private schools and six were state schools. In total, 1,137 students (81.4%) attended a state school and 261 (18.6%) attended a state-assisted private school.

Instruments

The instrument used was the School Attitude Assessment Survey-Revised (SAAS-R) developed by McCoach and Siegle (2003b), translated into Spanish using the technique of parallel back-translation (Brislin, 1986; Hambleton, 1994) and following the new guidelines provided by the International Test Commission

(Muñiz, Elosua, & Hambleton, 2013) together with the suggestions for implementing these proposed by Van der Vijver (2003). First, the original version of the SAAS-R was translated into Spanish by a native Spanish speaker with a degree in translation, majoring in English, and a lecturer in developmental and educational psychology with knowledge of both Spanish and English culture. Once completed, the Spanish translation was back-translated into English by a native English translator with a degree in Spanish and knowledge of both cultures. Then, the original version was compared with the back translation and translators made corrections to the final Spanish translation. The items obtained following this procedure were evaluated by a committee composed of three translators and three lecturers who were experts in the field of educational psychology. The committee selected those items that had retained their original meaning and devised the format and instructions of the scale in such a way that it remained identical to the original version, seeking to achieve satisfactory content validity of the test. Every effort was made to ensure consideration of linguistic, psychometric and cultural criteria in order to obtain a quality adaptation of the instrument (Hambleton & De Jong, 2003; Schweizer, 2010). No items were eliminated or significantly changed during the translation process. Table 1 shows the SAAS-R items translated into Spanish. The original items are available at McCoach and Siegle (2003b).

Cattell and Cattell's (1994) *g* factor test was employed to determine students' potential ability. Split-half reliability in the validation sample was .78.

Lastly, to determine actual academic performance (actual GPA), the arithmetic mean was calculated of all the marks obtained by students by the end of the school year for the totality of subjects studied.

Procedure

Prior to administering the tests, the necessary consent was sought from the competent authorities and school boards of the various schools. Once obtained, informed consent was then sought from the students' parents or legal guardians. The instrument was administered in the schools themselves in the second term of the 2011-2012 academic year during normal class hours. The tests were administered by collaborating researchers who had previously received instruction in the procedures to follow, laying particular emphasis on the voluntary nature of participation and the need for sincerity. On average, approximately 60 minutes were required to administer both tests.

Data analysis

The aim of the statistical analyses performed was to study the validity and reliability of the SAAS-R for a population of Spanish adolescent students. However, before performing the analyses, data distribution was examined to test for univariate and multivariate normality. Univariate normality showed skewness and kurtosis values below |1.5| for all items except those corresponding to the Goal Valuation factor (Items 14, 17, 19, 22, 23 and 32). In these cases, the indices of skewness and kurtosis exceeded the limits recommended, reaching values for skewness of between -1.7 and -3.1 and for kurtosis of between 2.9 and 10.6. Mardia's multivariate kurtosis coefficient was 592.64, indicating that the data were not multivariate normal.

For confirmatory factor analysis (CFA), the data collected using ordinal measures (such as Likert-type scales) could be considered categorical data. Even when the data appear to be approximately normally distributed, ordered categorical data are discrete and by definition cannot therefore be normally distributed (Finney & DiStefano, 2006). In addition, item distributions have a pronounced negative skew in most of the self-report scales of personal attributes such as personality, motivation and self-concept (Byrne, 2006).

Although it is assumed in categorical data analysis that each observed variable has an underlying scale that is both continuous and normally distributed, especially in scales with five or more categories (Byrne, 2006), problems begin to emerge as the observed item distributions diverge widely from a normal distribution. In particular, fit indices such as the CFI are underestimated when modeling non-normal ordered categorical data (Finney & DiStefano, 2006; Hutchinson & Olmos, 1998). This could lead to rejecting a correctly specified model.

When conducting structural equations modeling (SEM) with categorical data, the analysis must be based on the correct correlations, the polychoric correlation matrix. Particularly in the case of skewed responses, ordinary Pearson product-moment

Table 1
Spanish SAAS-R items

1. Mis clases son interesantes.
2. Soy inteligente.
3. Aprendo ideas nuevas rápidamente en el instituto.
4. Me alegro de venir a este instituto.
5. Este es un buen instituto.
6. Soy trabajador en el instituto.
7. Me llevo bien con mis profesores.
8. Me siento automotivado para hacer mis tareas.
9. Este instituto encaja bien conmigo.
10. El instituto es fácil para mí.
11. Me gustan mis profesores.
12. Mis profesores hacen que aprender sea interesante.
13. Mis profesores se preocupan por mí.
14. Ir bien en el instituto es importante para los objetivos de mi futura carrera.
15. Me gusta este instituto.
16. Comprendo conceptos complejos en el instituto.
17. Ir bien en el instituto es uno de mis objetivos.
18. Realizo las tareas con regularidad.
19. Es importante sacar buenas notas en el instituto.
20. Soy organizado con mis tareas.
21. Tengo varias estrategias para aprender cosas nuevas.
22. Quiero hacerlo lo mejor posible en el instituto.
23. Es importante para mí ir bien en el instituto.
24. Dedico mucho tiempo a mis tareas.
25. La mayoría de profesores de esta escuela son buenos profesores.
26. Soy un estudiante responsable.
27. Me esfuerzo mucho en mis tareas.
28. Me gustan mis clases.
29. Me concentro en mis tareas.
30. Reviso mis tareas antes de entregarlas.
31. Soy capaz de sacar sobresaliente en todo.
32. Quiero sacar buenas notas en el instituto.
33. Soy bueno aprendiendo cosas nuevas en el instituto.
34. Soy listo en el instituto.
35. Estoy orgulloso de este instituto.

correlations will tend to underestimate the "true" correlations and their functions, such as factor loading, and it is also possible that an erroneous lack of fit will be obtained for a true CFA model (MacCallum, Zhang, Preacher, & Rucker, 2002).

In an attempt to resolve these problems, two main approaches to modeling categorical data have been proposed (Bentler, 2005; Muthén & Muthén, 2004). The strategy suggested by Bentler (2005) and implemented in EQS 6.0 and subsequent versions, involves the use of an improved estimator of polychoric correlations together with a robust method such as maximum likelihood (ML) followed by robust computations based on an appropriate weight matrix and statistic such as the Satorra-Bentler (S-B) scaled χ^2 statistic. This yields optimal estimates and standard errors in sufficiently large samples (Bentler, 2005). When EQS analyses are based on categorical data, interpretation of model fit must be based on robust statistics (Bentler, 2005; Byrne, 2006).

CFA assessment of fit should be based on several indicators. One of the most commonly used fit indices is the chi-square statistic and its associated probability, which in the context of categorical and/or non-normal data is replaced by the Satorra-Bentler scaled chi-square, S-B χ^2. However, this test has the disadvantage of being sensitive to sample size. The comparative fit index (CFI) is one of the most widely used and reliable indices (Tanaka, 1993); values equal to or greater than .90 are considered to be indicative of acceptable model fit, although values equal to or greater than .95 are preferable, indicating a good model fit (Hu & Bentler, 1999). The root mean square error of approximation (RMSEA) is another of the most suitable indicators to examine model fit. Values below 0.06 are indicative of a close fit (Hu & Bentler, 1999).

Another strategy used to accommodate non-normal and/or categorical data is bootstrapping (Byrne, 2006; Finney & DiStefano, 2006; Nevitt & Hancock, 2001), a technique that substitutes reliance on theoretical χ^2 distribution and its assumptions. Bollen and Stine (1992) proposed the model-based bootstrap method, which is suitable for obtaining adjusted *p* values for the model fit statistics and fit indices, and evaluates overall model fit (Bentler, 2005). When data are non-normal, the Bollen-Stine model-based bootstrap method generally provides more accurate probability values than the S-B scaled χ^2 method (Nevitt & Hancock, 2001), although the bootstrap shows less power to identify misspecified models than the S-B χ^2 (Finney & DiStefano, 2006).

Related to the first objective, CFA was conducted to determine construct validity, based on a model with five first-order correlated factors, Model 1.

Related to the second objective, a model with five first-order factors and one second-order factor was estimated, Model 2. Because Model 1 and Model 2 are nested, their solutions were compared with the S-B χ^2 differences adjusted with the procedure proposed by Satorra and Bentler (2001), and directly compared with the Yuan-Bentler residual-based statistic (Yuan & Bentler, 1998).

These models were estimated using the maximum likelihood (ML) method, based on the polychoric correlation matrix, and tested with the S-B scaled χ^2 statistic (Satorra & Bentler, 2001) and the associated robust confirmatory fit indices provided by EQS 6.1 (Byrne, 2006). Furthermore, the analysis of Model 1 was performed using bootstrap ML estimation, based on the Bollen-Stine bootstrap *p* value and bootstrap adjusted χ^2 and goodness-of-fit indices provided by EQS (Bentler, 2005). In this second analysis, variables were treated as if they were continuous and Pearson correlation coefficients were computed.

Related to the third objective, several analyses were performed. First, three groups of subjects were identified, namely subjects whose performance was above (overachievers), equal to and below (underachievers) that expected, using the simple difference score method (Lau & Chan, 2001). In other words, a discrepancy score was calculated by subtracting the standardized performance score from the standardized ability score. Students whose discrepancy score was higher than +1 were selected as underachievers, while students with a discrepancy score below -1 were classified as overachievers; the comparison group of students in the equal to expectations range of achievement consisted of participants whose discrepancy scores were between +/- 1.

To test statistical differences among groups on the five factors of the SAAS-R, the General Linear Model (GLM) module of the PASW (version 20.0) was used. A multivariate analysis of variance (MANOVA) and a univariate analysis of variance (ANOVA) of repeated measures (Type III) were performed.

All statistical analysis used a confidence level equal to 95%.

Results

To test the internal structure of scores on the SAAS-R, confirmatory factor analysis based on the polychoric correlation matrix, together with the robust ML method followed by robust computations, was performed to test the fit to data of Model 1.

Once it had been confirmed that there were no offending parameter estimates (such as Heywood cases), model fit was assessed using absolute and incremental goodness of fit indices. Although the chi-square test indicated the model should be rejected (S-B $\chi^2_{(550)}$ = 1751.74, p<.001), the proposed model obtained a good absolute fit to the data. Given the influence of sample size on χ^2, other fit indexes are shown in Table 2.

Furthermore, given the significant multivariate non-normality of the data (normalized coefficient values for kurtosis = 217.85), the analyses were also performed using the ML model-based bootstrap method; the results obtained with 1000 bootstrap

Table 2
Fit indices

	S-B χ^2	df	p	CFI	NFI	NNFI	IFI	RMSEA	AIC	BIC[c]
Model 1[a]	1751.74	550	.000	.951	.931	.947	.951	.040	651.74	3332.51
Model 1[b]	796.11	550	.003	.990	.968	.989	.990	.017	–	–
Model 2	1897.59	555	.000	.945	.927	.936	.945	.044	672.41	3562.06

[a] = Model estimated with robust method and S-B χ^2 scaled statistic
[b] = Model estimated with the model-based bootstrap method
[c] = This index was estimated with AMOS-20

samples, of which 0 samples were unused, showed that the CFA estimates for Model 1 (Table 2) exhibited significant bootstrapped χ^2 values, with CFI and IFI above .95, and RMSEA below .05. With the exception of χ^2, which is influenced by sample size, the indices estimated using the robust method and those estimated with the model-based bootstrap method both showed a very satisfactory fit.

All the factor loadings obtained for each factor were statistically significant ($p<.001$), and all the standardized values were greater than .50 (Figure 1). The mean item loading for each scale was .83 for AS, .72 for ATT, .84 for ATS, .80 for GV, and .72 for M/S.

Intercorrelations among SAAS-R scales were moderate and significant ($p<.001$), ranging from .42 (ATS-M/S) to .70 (ATT-ATS), as shown in Figure 1.

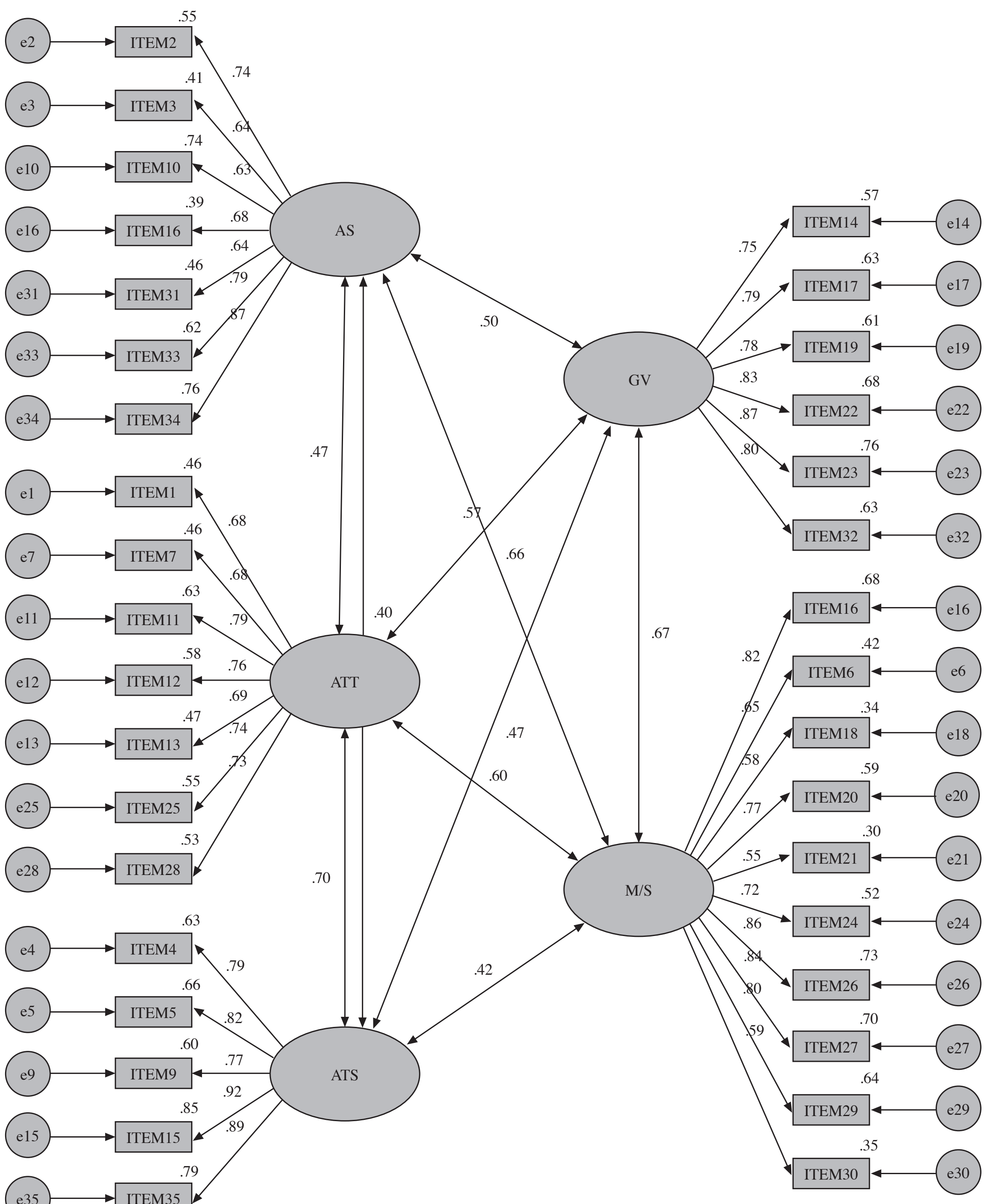

Figure 1. *CFA model of five first-order correlated factors. AS = Academic Self-Perception; ATT = Attitudes toward Teachers; ATS = Attitudes toward School; GV = Goal Valuation; M/S = Motivation/Self-Regulation*

Internal consistency reliability indices (Cronbach's alpha) were .86 for AS, .87 for ATT, .90 for ATS, .85 for GV, and .90 for M/S.

Related to the second objective, Models 1 and 2 were compared. The scaled-difference χ^2 test statistic (Satorra & Bentler, 2001) showed that the two models were not equivalent (S-B χ^2 dif. = 255.03, *df* = 5, *p* = .000), with the five first-order correlated factors fitting better to data. Direct comparison of residuals also showed a better fit for Model 1 (Yuan-Bentler residual test statistic = 13.8, *df* = 5, *p* = .02).

Group differences on SAAS-R scores were examined to provide concurrent validity evidence of instrument-criterion relationships, considering students with various levels of academic achievement and potential ability. Thus, the overachiever group consisted of a total of 228 students, the group whose performance was equal to their potential was composed of 933 students and the underachiever group comprised 237 students.

The SPSS System GLM Method I (Type III) was used to calculate the multivariate analysis of variance and analysis of variance tests. According to Wilks' criterion, the combined dependent variables (five SAAS-R scales) were significantly affected by achievement group, $F_{(10, 2784)} = 16.025$, $p<.0001$, $\eta_p^2 = .054$. Univariate tests for each SAAS-R scale reached statistical significance ($p<.01$), indicating that attitudes toward school and learning differed among groups. The effect size is low (<.20) in all cases following the interpretation of Cohen (1988).

To determine the nature of the group differences for each of the dependent variables, follow-up analyses were conducted using Games-Howell test. Compared with students whose performance was equal to or above their potential level, students who performed worse than expected obtained significantly lower scores on each SAAS-R scale. Comparing the scores of the overachiever and equal to potential groups, the former obtained significantly higher scores for Motivation/Self-Regulation and Academic Self-Perception (Table 3).

Discussion

Overall, the Spanish adaptation of the SAAS-R demonstrated good psychometric properties in a secondary student sample. Our findings suggest that the instrument appears to demonstrate evidence of acceptable construct validity, internal consistency reliability and criterion-related validity.

Regarding construct validity, confirmatory factor analysis supported the five-factor structure of the SAAS-R, and the fit indices indicated a good fit to data, in accordance with the criteria proposed by Hu and Bentler (1999). The indices obtained in this study by the correlated five-factor model also showed a slightly better fit with the data than those found in the initial validation study by McCoach and Siegle (2003b), and were similar to those reported in the study by Suldo et al. (2008). However, the latter study used item parcels in the CFA, a procedure which is not recommended because it can mask the true relationships among variables, leading to biased parameter estimates and fit indices (Bandalos, 2002; Finney & DiStefano, 2006). In addition, an examination of the factor loadings revealed that all items served as valid indicators of their corresponding theoretical constructs.

Table 3
Mean SAAS-R Scores by achievement group

Scale	Overachievers (n = 228) M	SD	Performance equal to potential (n = 933) M	SD	Underachievers (n = 237) M	SD	η_p^2
ASP	36.27_a	7.78	34.46_b	8.06	30.31_c	8.30	.048
ATT	36.20_a	8.01	35.73_a	8.07	32.88_c	8.48	.018
ATS	28.16_a	6.83	28.44_a	6.60	26.62_c	7.06	.010
GV	39.49_a	4.39	39.10_a	4.28	36.88_c	6.08	.034
M/S	56.14_a	9.80	51.70_b	10.65	44.46_c	12.30	.092

Note: Significant differences between group means are indicated by different letters. Means with the same subscript are not significantly different

The internal consistency reliability of the subscales and the total scale was also very acceptable, even more so than in other studies. Therefore, the results of this study, obtained in a more heterogeneous sample, support the previous findings of other studies with gifted students (McCoach & Siegle, 2003b) and gifted and non-gifted students (McCoach & Siegle, 2001; Suldo et al., 2008).

The pattern of intercorrelations among SAAS-R scales was very similar to that of previous validation studies. Thus, the highest correlation was between the scales Attitudes toward Teachers with Attitudes toward School (.70), as in the study by Suldo et al. (2008) who obtained .72. Meanwhile, the second highest correlation was between the scales Goal Valuation and Motivation/Self-Regulation (.67), while this figure was .74 in the study by McCoach and Siegle (2003) and .63 in that by Suldo et al. (2008). These moderate to high correlations between scales are troubling in terms of discriminant validity, but they are not unexpected theoretically. Relatively large correlations between subscales in turn suggest that they assess similar constructs, as also is evidenced for the high internal consistency of the total scale. Nevertheless, the results of model comparison yielded confirmation of the multidimensional structure of the instrument, favoring a model with five first-order correlated factors over a model with one second-order factor and five first-order factors.

The results of this study also evidenced criterion-related validity. The findings obtained using an objective indicator of academic performance demonstrated that academic underachievers had the lowest scores on each of the five subscales measured by the SAAS-R. Furthermore, this study shows that the SAAS-R is also capable of differentiating between students who perform above, below and equal to their potential. Most of the studies that have used the SAAS-R to examine differences between underachievers and non-underachievers have used samples of gifted students, whereas the present study encompassed the entire range of intellectual ability. The predominant subject of study in the USA has traditionally been the underachievement of high ability students (Colangelo, Kerr, Christensen, & Maxey, 2004; McCoach & Siegle, 2003b), whereas in China, the study of underachievers has been extended to the entire range of intellectual ability (Lau & Chan, 2001; Phillipson, 2008). In Spain, and even in Europe as a whole, little research has been conducted on underachievement, and therefore we believe that the present study may also help to arouse interest in the subject.

One limitation of this study could be the lack of a significant fit of the correlated five-factor model to the data, taking the chi-square exact test as the criterion. Although the model fit statistics indicated that the model met the criteria established by Hu & Bentler (1999), the chi-square exact test indicated that the model should be rejected. Although this is a frequent finding in CFA with several variables, and has also occurred in SAAS-R validation studies, it is also indicative of the existence of significant item cross loading.

A second study limitation concerns sample characteristics. Irrespective of the size and representativeness of the sample used in this study, which might equally be indicative of the study's strengths, the sample was not obtained from all parts of Spain. Nonetheless, there is nothing to suggest a priori that the sample used is not be representative of the population.

A third limitation is that measurement invariance is not given (Ziegler & Bensch, 2013), and this remains a question open to future research.

In sum, the SAAS-R is an instrument that has been shown to be valid and useful for the purposes for which it was designed, in a different linguistic and cultural context to that for which it was initially developed, with a large and representative sample of students. This instrument provides practitioners and researchers with a new tool for identifying secondary students who may be at risk of underachievement, and consequently to design programs to reverse underachievement.

Acknowledgements

This research was supported through research project GRE11-15, awarded by the Vice Chancellor for Research at the University of Alicante, and Research & Development Project No. EDU2012-32156, awarded by the Spanish Ministry of Economy and Competitiveness.

References

Bandalos, D.L. (2002). The effects of item parceling on goodness of fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*, 78-102.

Bentler, P.M. (2005). *EQS 6 structural equations program manual*. Encino, CA: Multivariate Software.

Bollen, K.A., & Stine, R.A. (1992). Bootstrapping goodness of fit measures in structural equation models. *Sociological Methods & Research*, *21*(2), 205-229.

Brislin, R.W. (1986). The wording and translation of research instruments. In W. L. Lonner & J. W. Berry (Eds.), *Field Methods in Cross-Cultural Research*. (pp. 137-164). Newbury Park, CA: Sage.

Byrne, B.M (2006). *Structural Equation Modeling with EQS. Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Cattell, R.B., & Cattell, A.K.S. (1994). Test de Factor «g» de Cattell [Test of g factor]. Institute for personality and ability testing. Spanish adaptation. Madrid, Spain: TEA Ediciones, S.A.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. New York: Academic Press.

Colangelo, N., Kerr, B., Christensen, P., & Maxey, J. (2004). A comparison of gifted underachievers and gifted high achievers. In S. Moon (Ed.), *Social/emotional issues, underachievement, and counseling of gifted and talented students* (pp. 119-132). Thousand Oaks, CA: Corwin Press.

Dowdall, C.B., & Colangelo, N. (1982). Underachieving gifted students: Review and implications. *Gifted Child Quarterly, 26*, 179-184.

Finney, S.J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G.R. Hancock & R.O. Mueller (Eds.), *Structural equation modeling: A* second course (pp. 269-314). Greenwich, CT: Information Age Publishing.

Hambleton, R.K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, *10*, 229-244.

Hambleton, R.K., & De Jong, J.H.A.L. (2003). Advances in translating and adapting educational and psychological test. *Languaje Testing, 20*(2), 127-134.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Hutchinson, S.R., & Olmos, A. (1998). Behavior of descriptive fit indices on confirmatory factor analysis using ordered categorical data. *Structural Equation Modeling: A Multidisciplinary Journal*, *5*, 344-364.

Kirk, C.M., Lewis, R.K., Scott, A., Wren, D., Nilsen, C., & Colvin, D.Q. (2012). Exploring the educational aspirations-expectation gap in eight grade students: Implications for educational interventions and school reform. *Educational Studies*, *38*(5), 507-519.

Lang, J.M., Watterman, J., & Baker, B.L. (2009). Computeen: A randomized trial of a preventive computer and psychological skills curriculum for at-risk adolescents. *Journal of Primary Prevent*, *30*, 587-603.

Lau, K.L., & Chan, D.W. (2001). Identification of underachievers in Hong Kong: Do different methods select different underachievers? *Educational Studies*, *27*(2), 187-200.

MacCallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*, 19-40.

McCoach, D.B. (2002). A validation study of the School Attitude Assessment Survey. *Measurement and Evaluation in Counseling and Development*, *35*, 66-77.

McCoach, D.B., & Siegle, D. (2001). A comparison of high achievers' and low achievers' attitudes, perceptions, and motivations. *Academic Exchange Quarterly, 5*, 71-76.

McCoach, D.B., & Siegle, D. (2003a). Factors that differentiate underachieving gifted students from high-achieving gifted students. *Gifted Child Quarterly, 47,* 144-154.

McCoach, D.B., & Siegle, D. (2003b). The School Attitude Assessment Survey-Revised: A new instrument to identify academically able students who underachieve. *Educational and Psychological Measurement, 63,* 414-429.

Muñiz, J., Elosua, P., & Hambleton, R.K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición [Guidelines for the translation and adaptation of test: Second Edition]. *Psicothema*, *25*(2), 151-157.

Muthén, L.K., & Muthén, B.O. (2004). *Mplus users's guide*. Los Angeles, CA: Authors.

Nevitt, J., & Hancock, G.R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling*, *8*(3), 353-377.

Phillipson, S.N. (2008). The optimal achievement model and underachievement in Hong Kong: An application of the Rasch model. *Psychology Science Quarterly, 50*(2), 147-172.

Pintrich, P.R., & DeGroot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40.

Rubenstein, L.D., Siegle, D., Reis, S.M., McCoach, D.B., & Burton, M.G. (2012). A complex quest: The development and research of underachievement interventions for gifted students. *Psychology in the Schools, 49*(7), 678-694.

Satorra, A., & Bentler, P.M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507-514.

Schweizer, K. (2010). The adaptation of assessment instruments to the various European languages. *European Journal of Psychological Assessment, 26*(2), 75-76.

Suldo, S.M., Shaffer, E.J., & Shaunessy, E. (2008). An independent investigation of the validity of the School Attitude Assessment Survey-Revised. *Journal of Psychoeducational Assessment, 26*(1), 69-82.

Tanaka, J.S. (1993). Multifacted conceptions of fit in structural equation models. In K.A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 10-39). Newbury Park, CA: Sage.

Van der Vijver, F.R.J. (2003). Test adaptation/translation methods. In R. Fernández-Ballesteros (Ed.), *Encyclopedia of psychological assessment* (pp. 960-964). Thousand Oaks, CA: Sage.

Yuan, K.-H., & Bentler, P.M. (1998). Normal theory based test statistics in structural equations modelling. *British Journal of Mathematical and Statistical Psychology, 51,* 289-309.

Ziegler, M., & Bensch, D. (2013). Lost in translation: Thoughts regarding the translation of existing psychological measures into other languages. *European Journal of Psychological Assessment*, *29*(2), 81-83.