European Journal of Investigation in Health, Psychology and Education
2016, Vol. 6, Nº 3 (Págs. 133-149)

e-ISSN 2254-9625 // www.ejihpe.es
doi: 10.1989/ejihpe.v6i3.175

Underachievement in Compulsory Secondary Education: a comparison of statistical methods for identification in Spain

Alejandro Veas, Raquel Gilar, Juan Luis Castejón and Pablo Miñano
University of Alicante (España)

This study compares the statistical methods employed for detecting underachievement, specifically the simple difference method, the regression method and the Rasch method. A sample of 1182 first- and second-year secondary students from 8 high schools in the province of Alicante participated in the study. The results showed a percentage of underachieving students that varies from 14.55% to 30.37%, depending on the statistical method employed. The Rasch method identified the highest number of underachieving students. Statistically significant differences were found between gender and type of student-underachieving and non-underachieving; however, no significant differences were detected between the course and type of student. This study confirms the importance of knowing the measurement properties of the statistical methods, how they affect the detection of underachieving students, and the main educational implications.

Keywords: Underachievement, simple difference method, regression method, Rasch method, Compulsory Secondary Education.

Rendimiento menor al esperado en Educación Secundaria Obligatoria: comparación de métodos estadísticos para su identificación en España. El presente estudio realiza una comparación de los métodos estadísticos más empleados en la detección del alumnado con rendimiento menor al esperado; concretamente el método de diferencias estandarizadas, el método de residuales de regresión y el método Rasch. Se empleó una muestra de 1182 alumnos de primer y segundo curso de Enseñanza Secundaria Obligatoria de 8 centros educativos de la provincia de Alicante. Los resultados muestran un porcentaje que oscila entre el 14.55% y el 30.37% de alumnos con rendimiento menor al esperado detectados en función del método estadístico empleado, siendo el método Rasch el que más alumnos identifica. Asimismo, se detectaron diferencias estadísticamente significativas entre el género y el tipo de alumnado (con rendimiento normal y con rendimiento menor al esperado), mientras que no se detectaron diferencias en función del curso y el tipo de alumnado. Este estudio confirma la importancia de conocer cómo afectan las propiedades de distintos métodos estadísticos en la detección del alumnado con rendimiento menor al esperado, así como sus principales implicaciones educativas.

Palabras clave: Rendimiento académico menor al esperado, método de diferencias estandarizadas, método de residuales de regresión, método Rasch, Educación Secundaria Obligatoria.

Correspondence: Alejandro Veas. Department of Development Psychology and Teaching. University of Alicante. Ctra. San Vicente del Raspeig, s/n. C.P.: 03080. Alicante (Spain). E-mail: alejandro.veas@ua.es

During the academic process of learning, students must face increasing levels of competence to complete curriculum objectives. However, some students, though they exhibit good levels of individual aptitude, may exhibit poor academic performance (Chan, 1999; Colangelo, Kerr, Christensen, & Maxey, 1993; McCall, Beach, & Lau, 2000). In this sense, the term underachievement has emerged as an important construct in the field of education during the last decades, and researchers have worked to detect and identify cognitive and non-cognitive variables which are involved (Lau & Chan, 2001; Matthews & McBee, 2007; McCoach & Siegle, 2003; Montgomery, 2003). The aims of the present study was to compare different statistical methods employed for detecting underachievement in high school students, and to analyze possible gender and course differences.

Definition of underachievement

First, no consensual definition of underachievement has been accepted by the scientific community (McCoach & Siegle, 2011). In scientific literature, there is a general agreement that underachievement is a discrepancy between what can be expected and what is actually achieved (Phillipson, 2008).Researchers have addressed underachievement in a variety of contexts, including studies related to the operationalization of the concept (Ziegler, Ziegler, & Stoeger, 2012), the possible inclusion of students with learning disabilities into the underachievement framework (Fletcher, Denton, & Francis, 2005) or the analysis of underachieving students with emotional and behavioral disorders (Lane, Gresham, & O'Shaughnessy, 2002).

Clearly, underachievement is a multidimensional construct that involves different variables. Analyses of these variables have focused on underachieving gifted students (Chan, 1999; Dixon, Craven, & Martin, 2006; Obergriesser & Stoeger, 2015; Ziegler & Stoeger, 2003), especially in the United States (Figg, Rogers, McCormick, & Low, 2012; McCoach & Siegle, 2003; Reis & McCoach, 2000; Reis & Greene, 2002). However, the authors of this study, in agreement with Dittrich (2014), support the assumption that underachievement is not reserved exclusively for gifted students but to all of the students situated in different intelligence levels who may also be influenced by personality factors, family-related factors and school-related factors. Indeed, the treatment of these factors through educational interventions could lead to a better self-concept and academic achievement (Álvarez, Suárez, Tuero, Núñez, & Valle, 2015; Valle et al., 2015, 2015; Veas, Castejón, Gilar, & Miñano, 2015).

Underachievement in Spain

In comparison with other countries, few studies focus on the study of underachievement in Spain, and the majority of studies are related to gifted students. One of the most important studies was developed in Madrid by García-Alcañiz (1991)

where the percentage of gifted students with school failure or dropout was 30%, similar to normal the population. Jiménez and Álvarez (1997) confirmed the same percentage of students with high IQ and low achievement during the first school years. Broc (2010) treated underachievement in the context of school failure and absenteeism and formulated a theoretical model to explore the reasons for low academic achievement in students with high academic potential. Lastly, and according to the more recent input from the scientific literature, Veas, Castejón, Gilar, & Miñano (2016) detected a total of 181 (28.14% in a sample of 643) underachieving, Spanish first-year high school students using the Rasch model.

Because few studies explore underachievement in Spain, there is no evidence of specific cultural factors that affect the occurrence of underachievement (Reis & McCoach, 2000). Therefore, to better understand the phenomena, it is necessary to determine if the percentage of underachieving students, identified with different methods, is similar in a different cultural context such as Spain. Furthermore, it is also important to analyze some of these factors related with underachievement, such as gender and course (Driessen & van Langen, 2013).

Statistical methods for detecting underachievement

Before any type of educational intervention or the analysis of the variables involved in the underachievement process, underachieving students must be identified. From a methodological perspective, the traditional statistical methods are the absolute split method, the simple difference method and the regression method (Lau & Chan, 2001). When using the absolute split method, the researcher uses an arbitrary limit on the top of the mental ability and the bottom of academic performance after the conversion of punctuations to standard scores. This method has been used specifically in studies on gifted underachieving students (Peterson & Colangelo, 1996; Vlahovic-Stetic, Vidovic, & Arambasic, 1999).

The simple difference method is based on the discrepancy between the standardized performance score and the standardized ability score. When the difference is based on an arbitrary limit (normally 1 standard deviation), a student could be regarded as underachieving (D<-1) or overachieving (D>1). According to Lau and Chan (2001), this method more appropriately identifies underachievement at all levels of ability.

The third method is the regression method, which is based on the deviation of the students´ score from the regression line of the achievement measure on the ability measure. Students are considered as underachieving if this deviation is negative and greater than one standard error of estimate. While this method seems to have better reliability than the method of simple difference scores, it generates a constant proportion

of underachieving students (McCoach & Siegle, 2011; Plewis, 1991; Ziegler et al., 2012).

The statistical methods described above are based on the use of arbitrary cutoff and the use of standardized transformations. According to Phillipson & Tse (2007), this type of comparison does not suppose the assumption that the original data are interval in nature. To improve the objective use of the interval scale, the last method used to identify underachievement is the Rasch model (Phillipson, 2008; Phillipson & Tse, 2007). This model is one of the most well-known among item response theories, representing the variability of a construct based on the calibration of ordinal data from a shared measurement scale. The Rasch model establishes that the difficulty of the items and the ability of the subjects can be measured on the same scale and that the likelihood that a subject responds correctly to an item is based on the difference between the ability of the subject and the difficulty of the item (Rasch, 1960/1980; Wright & Stone, 1979). Both ability and difficulty are estimated using logit units because a logarithmic scale is used. The main advantage of the logarithmic scale is the establishment of homogeneous intervals through the range of variables, which means that the same difference between the difficulty parameter of an item and the ability of a subject involves the same probability of success along the entire scale (Preece, 2002). The adjustment of this interaction can be conducted using residual measures and can be standardized for a particular item or subject in two ways (Bond & Fox, 2007). On one side is Outfit, based on the sum of squared standardized residuals of every item encountered by person n divided by the number of items to which person n responded. On the other side is Infit, a measure that eliminates the extreme scores that influence the Outfit by using the residuals of individuals whose ability levels are in the closest range to a particular item. Both indexes are indicated as the mean squares in the form of chi-square statistics divided by their degrees of freedom, which imply a ratio scale form with a range from 0 to positive infinitive. Therefore, values below 1 indicate a higher than expected fit of the model, and values greater than 1 indicate a poor fit of the model.

With Phillipson and Tse´s (2007) Rasch model, two validated tests are used for measuring ability and achievement. However, it is important to highlight the possibility to calibrate the General Points Average (GPAs) of students as the main measure of underachievement in Spanish schools (Veas et al., 2016). The analysis of the conceptual and methodological processes in comparing school grades have been studied extensively in the last quarter of the twentieth century, especially in the United Kingdom (Fitz-Gibbon, Vincent, & Britain, 1994; Forrest & Vickerman, 1982; Goldstein & Cresswell, 1996; Goldstein & Thomas, 1996). The authors of the present study consider the inter-subject comparability approach as an appropriate model in which the influence of the difficulty level of the subjects and the proficiency level of the students can be adjusted according to the Rasch´s parameters. This approach has been tested – with some

variation in the procedures - in different countries with positive results (Coe, 2007, 2008; Korobko, Glas, Bosker, & Luyten, 2008; TQA, 2006, 2007).

The present study

Given the different characteristics of the statistical methods (Lau & Chan, 2001; Phillipson, 2008; Phillipson & Tse, 2007), it is important to make comparisons to highlight the possible levels of association among them and the possible variations of the capacity of detection along the entire capacity continuum. Therefore, and as an extension of the study of Veas et al. (2016), the hypotheses of the present study were:

(1) The simple difference method, the regression method, and the Rasch method significantly identify a different number of underachieving students. (2) All of the three statistical methods detect a higher number of underachieving boys than underachieving girls. (3) All of the three statistical methods detect a similar proportion of underachieving students both in the first and the second course of ESO.

METHOD

Participants

Using the school as the sampling unit and taking into account geographical areas from the province of Alicante, random cluster sampling was used, selecting 8 high schools in the province of Alicante. A total of 1229 students in the first and second years of Compulsory Secondary Education participated in the study. Of these, 47 (3.82%) were excluded due to coding errors or a lack of qualifications because they had special education needs or because they did not have parental consent, resulting in a total of 1182. Six hundred and nineteen students were enrolled in the first course; whereas, 563 were enrolled in the second course. Overall, 53.29% of the students were male, and 46.71% were females. Childhood socioeconomic status (SES) was indexed according to parental occupation. There was a wide range of socioeconomic status with a predominance of middle class children. This classification was based on the level of incomes and the level of studies of the families. The regional education counselors determined the childhool socioeconomic statuses (SES) through a questionnaire registered with the responses of the students. The variables used were parents´ professions, professional situation and level of studies, number of books at home, cultural and sporting activities and availability of technological means at home. The Chi-square test was used to determine differences between the gender of the sample and the gender of the national student population (51.3% boys and 48.7% girls), supporting the absence of gender differences between the sample and the population (χ^2=0.28, *df*=1, *p*>.05).

Measures

Academic performance: Numerical GPAs from 9 mandatory courses, which the schools provided at the end of the school year, were considered. The courses recorded were Spanish Language and Literature, Natural Sciences, Valencian Language, Social Sciences, Mathematics, English, Technology, Art Education, and Physical Education. Students' scores showed high reliability with a Cronbach´s alpha of 0.93.

Cognitive ability: Students´ scholar ability was estimated using the Battery of Differential and General Skills (Yuste, Martínez, & Gálvez, 2005). This Spanish battery measures the capacities and academic abilities of students. There are six subscales: Analogies (A), Series (S), Matrices (M), Complete (C), Problems (P), and Figures Fit (E). Each subscale is measured with 32 items with five response options for which only one option is correct, producing a total of 192 items. For this study, Cronbach´s alpha values for each subscale were .83, .89, .79, .83, .77, and .87, respectively. Furthermore, a general intelligence quotient (IQ) could be obtained based on the punctuations from the distinct differential skills. The Cronbach´s alpha of the total IQ was .83.

Procedure

Prior to data collection, the necessary permission was requested from the educational administration and school boards of the various schools. After obtaining these permissions, the parents or legal guardians of the students had to provide the corresponding informed consent. Data collection was performed in the schools themselves during the second trimester of the school year and during normal school hours. The data were collected by collaborating researchers previously trained in the standards and guidelines for data collection.

Data analysis

For the identification of underachieving students in all capacity continuums, the simple difference method, the regression method and the Rasch model were employed. First, the simple difference method was calculated based on a punctuation of the discrepancy between the standardized performance score and the standardized ability score. The students whose punctuation of discrepancy was lower than -1 were identified as underachieving. Second, the regression method was calculated, employing the total IQ from BADyG as the predictor and the average grade of each student as the criteria. Students with a residual punctuation lower than -1 were identified as underachieving. SPSS version 21.0 software was used for both methods.

Lastly, for the use of the Rasch method, BADyG and GPAs were analyzed using Winsteps version 3.81 statistical software (Linacre, 2011) for which estimates were based on the joint maximum likelihood (Bond, 2003; Linacre, 2012). BADyG was calibrated with the dichotomous Rasch model, whereas GPAs were calibrated with the

Partial Credit Model (PCM) (Wright & Masters, 1982). Once fit indices from both measures were observed, the Rasch model allowed for testing the hypothesis that two tests measure the same underlying construct (Bond & Fox, 2007). This comparison was tested by elaborating a scatter plot of students. Rasch responses to both tests observing whether the points lie between 95% confidence bands (Phillipson, 2008). Those points outside the 95% confidence bands indicated that the achievement level was not what was expected. Once the different methods were implemented, they were compared by using the significant chi-square and the Phi coefficient (Lau & Chan, 2001), which indicate the levels of association between the methods employed and the proportion of students identified as underachieving and non-underachieving.

RESULTS

The exploratory analysis of the data shows that all of the variables followed a normal distribution, with values of skewness and kurtosis between +/-1. The mean of the BADyG IQ was 100.6 (*SD*=15.8) with a range of punctuations between 58 and 150. The mean of the final achievement was 6.3 (*SD*=1.8) and varied between 1.44 and 10.

Previous to the identification of underachieving students with the Rasch method, the analysis of the fit of the grades was conducted based on the inter-subject comparability approach. We used an approximate range of 0.8 to 1.2 for Infit and Outfit (Bond & Fox, 2007). Although not shown, the first fit values indicated a lack of fit in the majority of subjects, so recoding scores (Korobko et al., 2008) was performed using values that were based on the qualitative scores in Spanish schools (poor, sufficient, good, notable and outstanding). Therefore, the values employed were: 1 for categories 1, 2, 3, and 4; 2 for categories 5 and 6 (sufficient and good); 3 for categories 7 and 8 (notable); and 4 for categories 9 and 10 (outstanding). The new calibration of the courses provided a good fit for the data except for physical education (Infit=1.40; Outfit=1.54). The analysis of Differential Item Functioning (DIF) estimated the distribution of the difficulty parameter in the sample of boys and girls. The results show that the subject Visual Arts Education is easier for girls, and the difference is statistically significant (Mantel χ^2=23.518; $p\leq.00$). No differences were found in the rest of the subjects. Therefore, both Psychical Education and Visual Arts Education were eliminated, according to the requirements of the Rasch model, which imply that the data must fit the model to be accepted (Bond & Fox, 2001).

For the analysis of unidimensionality, a principal component analysis of the residual score was conducted (Linacre, 1998). The results showed a principal factor that was able to explain 69.3% of the variance of the latent trait with a wide difference between the weight of the first factor and the next (Eigenvalue=1.4), which favors the unidimensionality of the model.

With respect to the Rasch calibration of the BADyG, each block was analyzed separately. The item analysis demonstrated that the majority of items fit the model satisfactorily with values within 0.80 and 1.20. Regarding person fit, approximately 95% of students fit the Rasch model (Bond & Fox, 2001, pp. 176-177, Phillipson & Tse, 2007).

In the first course of secondary education, 82, 111, and 179 students were identified as underachieving with the simple difference method, the regression method and the Rasch method, respectively; whereas, 71, 90, and 180 students were detected in the second year of secondary education, following the same order of statistical methods. The analysis made with the Rasch method was produced after adjusting the school grade scores and BADyG scores to align with a mean of 0 and *SD* 1 (Bond & Fox, 2001, p.57). The scatterplot of person logit school grades scores and person logit BADyG scores was produced for each course (Figure1 and Figure 2).

Figure 1. Person logit school grades plotted against person logit Badyg scores in the first course of ESO, with 95% confidence bands

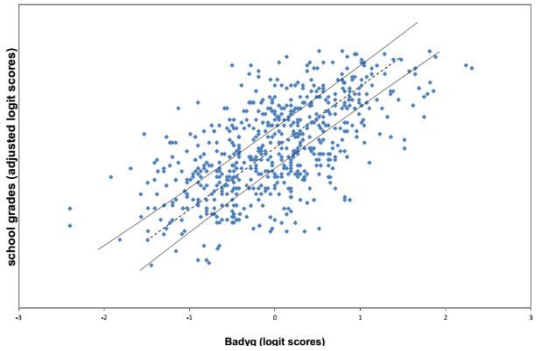
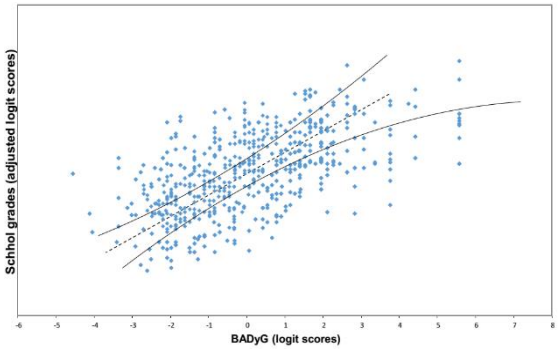


Figure 2. Person logit school grades plotted against person logit Badyg scores in the second course of ESO, with 95% confidence bands



Two-by-two tables were created to use a cross-tabulation procedure for comparing the statistical methods in each course. Chi-square tests and Phi coefficients were obtained to examine the relationship of each pair of selection methods and the proportion of students identified as underachieving and non-underachieving (Table 1 and Table 2).

In table 1, comparison between the Rasch method and the simple difference method showed 63 underachieving students selected from both methods, which represents 76.8% of the total of underachieving selected by the simple difference method and 35.2% of the total of underachieving selected by the Rasch method.

Therefore, the difference in the number of underachieving students detected by each method is considerably high, according to the significant values of chi-square and the Phi coefficient (χ^2=105.55, $p\leq.001$; φ=.413, $p\leq.001$). When comparing the Rasch method with the regression method, the number of underachieving detected by both methods is higher, although there are only 81 students identified as underachieving with the Rasch method; whereas only 13 students were identified with the regression method. The relationship of this pair of selection methods is statistically significant (χ^2=231.93, $p\leq.001$; φ=.612, $p\leq.001$). Lastly, comparison between the regression method and the simple difference method shows a higher percentage of underachieving students detected by both methods as 79% of the total of underachieving students were selected by the simple difference method and 58.6% of the total of underachieving selected by the regression method. Again, the level of association between this pair of methods and the percentage of underachieving and non-underachieving students is significant (χ^2=41.63, $p\leq.001$; φ=.625, $p\leq.001$.

Table 1. Comparison between pairs of the three statistical methods in detecting underachieving students of first year of ESO

		Non-underachieving	Underachieving	Total	Chi-square	Phi
		Rasch method				
Simple difference method	Non-underachieving	421	116	537		
	Underachieving	19	63	82		
Total		440	179	619	105.55*	.413*
Regression method	Non-underachieving	427	81	508		
	Underachieving	13	98	111		
Total		440	179	619	231.93*	.612*
		Regression method				
Simple difference method	Non-underachieving	491	46	537		
	Underachieving	17	65	82		
Total		508	111	619	241.63*	.625*

*$p\leq.01$

Table 2. Comparison between pairs of the three statistical methods in detecting underachieving students of second year of ESO

		Non-underachieving	Underachieving	Total	Chi-square	Phi
		Rasch method				
Simple difference method	Non-underachieving	375	117	492		
	Underachieving	8	63	71		
Total		383	180	563	120.35*	.503*
Regression method	Non-underachieving	381	92	473		
	Underachieving	2	88	90		
Total		383	180	563	213.28*	.642*
		Regression method				
Simple difference method	Non-underachieving	458	34	492		
	Underachieving	15	56	71		
Total		473	90	563	239.24*	.911*

*$p \leq .01$

Tables 3 and table 4 show the analysis of gender and course differences in each statistical method. In table 3, gender differences of underachieving students are detected in each of the methods with significant chi-square values and Phi coefficients. According to the analysis, more boys than girls are designated as underachieving, representing more than 50% of the total underachieving sample for each method. Finally, in table 4, there is a similar number of underachieving students in each of the courses and in all of the methods employed with no significant chi-square values and Phi coefficients. Clearly the largest number of underachieving students was detected with the Rasch model, with a total of 30.37% of the total sample.

Table 3. Gender analysis of the three statistical methods in selecting underachieving

		Gender		Total	Chi-square	Phi
		Boys	Girls			
Simple difference method	Non-underachieving	512	498	1010		
	Underachieving	118	54	172		
Total		630	552	1182	18.93*	.127*
Regression method	Non-underachieving	502	498	1000		
	Underachieving	128	54	182		
Total		630	552	1182	25.06*	.146*
Rasch method	Non-underachieving	394	429	823		
	Underachieving	236	123	359		
Total		630	552	1182	32.04*	.165*

*$p \leq .01$

Table 4. Course analysis of the three statistical methods in selecting underachieving

		Course		Total	Chi-square	Phi
		First	Second			
Simple difference method	Non-underachieving	537	473	1010		
	Underachieving	82	90	172		
Total		619	563	1182	1.77	-.039
Regression method	Non-underachieving	508	492	1000		
	Underachieving	111	71	182		
Total		619	563	1182	6.40	.074
Rasch method	Non-underachieving	440	383	823		
	Underachieving	179	180	359		
Total		619	563	1182	1.30	-.039

*$p \leq .01$

DISCUSSION AND CONCLUSIONS

The aims of the present study was to compare the simple difference method, the regression method, and the Rasch model in detecting underachievement; and to analyze gender and course differences between underachieving and non-underachieving students with each statistical method. For first objective, major level of detection was observed when using the Rasch model with 30.37% of underachieving students identified in the total sample of first and second year of Compulsory Secondary Education. The simple difference method detected 14.55% of underachieving students; whereas, the regression method detected 15.39%. The last two methods have similar percentages of underachieving detection, and many of these students were identified as underachieving by both methods. In contrast, the Rasch model detected 87 more underachieving students than the simple difference method and 68 more than the regression method.

This variation between the Rasch method and the other two methods is explained by measurement properties. Both the simple difference method and the regression method are highly dependent on sample parameters (Phillipson, 2008). For instance, when using the simple difference method, a student would need more levels of discrepancy between achievement and ability. When using the regression method, a large deviation of the prediction of achievement on ability is needed to identify underachievement; therefore, those students with high intellectual ability will have more probabilities to be chosen (Lau & Chan, 2001). Fletcher, Denton and Francis (2005) identify several problems with these methods, including problems of test reliability, the assumption of normality in measures and the use of discrepancy scores as the basis for the classification of students as underachieving. In contrast, the measurement properties of the Rasch model overcome previous limitations by converting ordinal data to linear measures based on a logarithmic scale, which is non-dependent sample. The use of a log-odds unit implies an interval scale in which differences between logits are homogeneous (Bond & Fox, 2001) and the calibration of both measures (ability and achievement) provides objective measures of persons and items in the same logit scale. Based on this method, it is possible to compare grades under the subject comparability approach (Coe, 2008; Newton, 2005). Reducing the number of categories for all courses, eliminating Physical Education to obtain adequate levels of fit (Wright, Linacre, Gustafson, & Martin-Lof, 1994) and eliminating the Arts and Visual Education subject because it had a significant Differential Item Functioning was necessary. The subjects analyzed together aim to measure overall academic performance, show good values of factor loadings in the principal component analysis, and confirm the unidimensionality of the construct. Because comparing different procedures for measuring academic achievement

is important, future studies may need to use achievement tests to contrast the quality of the use of grades in Spanish schools when detecting underachieving students.

Regarding the second objective, important gender differences were observed between non-underachieving and underachieving students with the total sample. Though nearly the same proportion of boys and girls are included in the non-underachieving group, a higher proportion of boys are identified as underachieving in comparison with girls, specifically 68% with the simple difference method, 70.32% in the regression method, and 65% with the Rasch method. During the last decades, studies that have focused on the relationship between gender and academic achievement have highlighted diverse results. There is a common thought that the trend of male underachievement has been evident for at least the last decade (Driessen & van Langen, 2013; Gibb, Fergusson, & Horwood, 2008). Our results confirm previous studies that showed that significant differences were in favor of girls (Eurydice, 2010). Girls appear to have established themselves as more reliable in terms of passing grades than their male peers. This situation began by the mid-1990s as boys began to emerge as significantly less successful than girls in terms of learning outcomes. There could be many influences at this point, as other related variables such as socio-economical level or disadvantaged backgrounds clearly affect this relationship. In this sense, it is also important to mention the stereotypical views of gender related to abilities. Literature reviews have identified this tendency, as well. For example, Hyde & Linn (2014) concluded in a meta-analysis that there were more similarities than differences between boys and girls, even in those areas such as mathematics or science where typical gaps have been detected. Therefore, the analysis of contextual factors is quite relevant to detect direct or indirect relations between gender and underachievement.

With respect to the analysis of the frequencies of underachieving students in the first and the second course, it is clear that no significant and significant differences have been found. That they were consecutive courses could be a possible explanation. The main reason for this comparison is that important changes happen in Secondary Education (Eccles & Roeser, 2011), e.g., new and bigger facilities, more professors for each subject, new partners, etc., which could imply a difficult process for some students in our cultural context (Pérez & Castejón, 2008).

At this point, it is important to highlight some limitations of the present study. First, we referred to a global underachievement instead of an underachievement index in a specific area, which implies a major probability of obtaining a higher number of underachieving students. Second, for a more objective measure of the subjects, it would be advisable to reduce the number of grades for evaluation, especially in the lowest categories as they are assigned to a very low proportion of students. Third, this study does not analyze the transition from primary to secondary education, or the transition from the first level to highest level of secondary education. For this reason, future studies

should be made with Primary students, as well as longitudinal studies in Secondary Education, in order to observe the whether the dynamic process of underachievement declines during the adolescence.

To conclude, this study highlights the need to revise and compare different statistical methods used to detecting underachieving students and the detection of differences in some important variables such as gender and course related to academic achievement. All of the statistical methods show an important percentage of these students in the first and second year of Compulsory Secondary Education. The Rasch model identified the most number of underachieving students, confirming the limitations of the other methods based on the use of cut-off points and the possibility that underachieving subgroups exist across the ability levels, especially in the medium and low scale of the continuum, according to the scientific literature (Reis & McCoach, 2000; Ritchotte, Matthews, & Flowers, 2014; Snyder & Linnenbrink-Garcia, 2013). Because it is possible to provide individual detection of underachievement, it would be necessary to develop educational programs adapted to the cultural factors in Spain and the possible underachieving subgroups and to analyze these differences using cognitive, motivational and contextual variables (Baker, Bridger, & Evans, 1998; Chan, 1999; McCoach & Siegle, 2003; Obergriesser & Stoeger, 2015).

Funding

The present work was supported by the Spanish Ministry of Economy and Competitiveness (Award number: EDU2012-32156) and the Vice Chancelor for Research of the University of Alicante (Award number: GRE11-15). The corresponding author is funded by the Spanish Ministry of Economy and Competitiveness (Reference of the grant: BES-2013-064331).

REFERENCES

Álvarez, A., Suárez, N., Tuero, E., Núñez, J.C., Valle, A., & Regueiro, B. (2015). Implicación familiar, autoconcepto del adolescente y rendimiento académico. *European Journal of Investigation in Health, Psychology and Education, 5*(3), 293-311. doi: 10.1989/ejihpe.v5i3.133

Baker, J.A., Bridger, R., & Evans, K. (1998). Models of underachievement among gifted preadolescents: The role of personal, family, and school factors. *Gifted Child Quarterly, 42*(1), 5-15. doi: 10.1080/02783190609554366

Bond, T.G. (2003). Validity and assessment: a Rasch measurement perspective. *Metodología de las Ciencias del Comportamiento, 5*, 179-194.

Bond, T.G., & Fox, C.M. (2001). *Applying the Rasch model: Fundamental measurement in the human science.* Mahwah, NJ: Erlbaum Associates.

Bond, T.G., & Fox, C.M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences.* New York, NY: Psychology Press.

Broc, M.A. (2010). *Un estudio de meta-análisis sobre las causas del absentismo escolar y propuestas de prevención e intervención para la educación primaria y secundaria obligatoria. Memoria de Investigación.* Zaragoza: Departamento de Educación, Cultura y Deporte del Gobierno de Aragón, España.

Chan, D.W. (1999). Reversing underachievement: Can we tap unfulfilled talents in Hong Kong? *Educational Research Journal, 14*(2), 177-190.

Coe, R. (2007). Common examinee methods. In P. Newton, J. Baird, H. Goldstein, H. Patrick, and P. Tymms (Ed.), *Techniques for monitoring the comparability of examination standards* (pp. 331-367). London: QCA.

Coe, R. (2008). Comparability of GCSE examinations in different subjects: An application of the Rasch model. *Oxford Review of Education, 34*(5), 609-636. doi: 10.1080/03054980801970312

Colangelo, N., Kerr, B., Christensen, P., & Maxey, J. (1993). A comparison of gifted underachievers and gifted high achievers. *Gifted Child Quarterly, 37*(4), 155-160.

Dittrich, E. (2014). Underachievement leading to downgrading at the highest level of secondary education in the netherlands: A longitudinal case study. *Roeper Review, 36*(2), 104-113. doi: 10.1080/02783193.2014.884201

Dixon, R.M., Craven, R., & Martin, A. (2006). Underachievement in a whole city cohort of academically gifted children: What does it look like? *Australasian Journal of Gifted Education, 15*(2), 9-15.

Driessen, G., & van Langen, A. (2013). Gender differences in primary and secondary education: are girls really outperformance boys? *International Review of Education, 59*(1), 67-86.

Eccles, J.S., & Roeser, R.W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence, 21*(1), 225-241. doi: 10.1111/j.1532-7795.2010.00725.x

Eurydice (2010). *Gender differences in educational outcomes*. European Commision. Retrieved from http://eacea.ec.europea.eu/education/eurydice/documents/thematic_reports/120EN.pdf

Figg, S.D., Rogers, K.B., McCormick, J., & Low, R. (2012). Differentiating low performance of the gifted learner achieving, underachieving, and selective consuming students. *Journal of Advanced Academics, 23*(1), 53-71. doi: 10.1177/1932202X11430000

Fitz-Gibbon, C.T., Vincent, L., & Britain, G. (1994). *Candidates' performance in public examinations in mathematics and science.* London, UK: School Curriculum and Assessment Authority.

Fletcher, J.M., Denton, C., & Francis, D.J. (2005). Validity of alternative approaches for the identification of learning disabilities: Operationalizing unexpected underachievement. *Journal of Learning Disabilities, 38*(6), 545-552. doi: 10.1177/00222194050380061101

Forrest, G.M., & Vickerman, C. (1982). *Standards in GCE: Subject pairs comparisons, 1972-80, occasional publication 39*. Manchester, UK: Joint Matriculation Board.

García-Alcañiz, E. (november, 1991). Diferencias intelectuales y de rendimiento académico entre chicos-chicas bien dotados y la media. Paper presented at the *I International Congress of Psychology and Education.* Madrid, Spain.

Gibb, S.J., Fergusson, D.M., & Horwood, L.J. (2008). Gender differences in educational achievement to age 25. *Australian Journal of Education, 52*(1), 63-80. doi: 10.1177/000494410805200105

Goldstein, H., & Cresswell, M. (1996). The comparability of different subjects in public examinations: A theoretical and practical critique. *Oxford Review of Education, 22*(4), 435-442. doi: 10.1080/0305498960220404

Goldstein, H., & Thomas, S. (1996). Using examination results as indicators of school and college performance. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 159*(1), 149-163. doi: 10.2307/2983475

Hyde, J.S. (2014). Gender similarities and differences. *Annual Review of Psychology, 65*, 373-398. doi: 10.1146/annurev-psych-010213-115057

Jiménez, C., & Álvarez, B. (1997). Alumnos de Alta Capacidad y Rendimiento Escolar Insatisfactorio. *Revista de Educación, 313*, 279-295.

Korobko, O.B., Glas, C.A., Bosker, R.J., & Luyten, J.W. (2008). Comparing the difficulty of examination subjects with item response theory. *Journal of Educational Measurement, 45*(2), 139-157. doi: 10.1111/j.1745-3984.2007.00057.x

Lane, K.L., Gresham, F.M., & O'Shaughnessy, T.E. (2002). Serving students with or at-risk for emotional and behavior disorders: Future challenges. *Education and Treatment of Children, 25(4)*, 507-521.

Lau, K., & Chan, D.W. (2001). Identification of underachievers in hong kong: Do different methods select different underachievers? *Educational Studies, 27*(2), 187-200. doi: 10.1080/03055690120050419

Linacre, J.M. (1998). Structure in Rasch residuals: Why principal component analysis? *Rasch Measurement Transactions, 12*, 636.

Linacre, J. (2011). *Winsteps Rasch measurement computer program [computer sofware]*. Chicago: Winsteps.

Linacre, J. (2012). *A User´s Guide to Winsteps & Ministep Rasch-Model Computer Programs*. Program Manual 3.74.0. 2012. Available online at: http://www.winsteps.com/winman

Matthews, M.S., & McBee, M.T. (2007). School factors and the underachievement of gifted students in a talent search summer program. *Gifted Child Quarterly, 51*(2), 167-181. doi: 10.1177/0016986207299473

McCoach, D.B., & Siegle, D. (2003). Factors that differentiate underachieving gifted students from high-achieving gifted students. *Gifted Child Quarterly, 47*(2), 144-154. doi: 10.1177/001698620304700205

McCoach, D.B., & Siegle, D. (2011). Underachievers. In R.J. Levesque (Ed.), *Encyclopedia of Adolescence* (pp. 3025-3032). New York, NY: Springuer Science & Business Media.

McCall, R.B., Beach, S.R., & Lau, S. (2000). The nature and correlates of underachievment among Elementary School children in Hong Kong. *Child Development, 71*(3), 785-801. doi: 10.1111/1467-8624.00185

Montgomery, D. (2003). Inclusive education for able underachievers: Changing teaching for learning. In D. Montgomery (Ed.), *Able underachievers* (pp. 127-149). London, England: Whurr.

Newton, P.E. (2005). Examination standards and the limits of linking. *Assessment in Education, 12*(2), 105-123. doi: 10.1080/09695940500143795

Obergriesser, S., & Stoeger, H. (2015). The role of emotions, motivation, and learning behavior in underachievement and results of an intervention. *High Ability Studies, 26*(1), 167-190. doi: 10.1080/13598139.2015.1043003

Pérez, A.M., & Castejón, J.L. (2008). Changes in cognitive-motivational factors resulting from the process of school transition. In A. Valle et al., (Eds). *Handbook of instructional resources and their applications in the classroom* (pp.223-237). New York, NY: Nova Science Publishers.

Peterson, J.S., & Colangelo, N. (1996). Gifted achievers and underachievers: A comparison of patterns found in school files. *Journal of Counseling & Development, 74*(4), 399-407.

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

Phillipson, S.N., & Tse, A.K. (2007). Discovering patterns of achievement in Hong Kong students: An application of the Rasch measurement model. *High Ability Studies, 18*(2), 173-190. doi: 10.1080/13598130701709640

Plewis, I. (1991). Underachievement: A case of conceptual confusion. *British Educational Research Journal, 17*(4), 377-385. doi: 10.1080/0141192910170407

Preece, P. (2002). Equal-interval measurement: The foundation of quantitative educational research. *Research Papers in Education Policy and Practice, 17*(4), 363-372. doi: 10.1080/0267152022000031397

Rasch, G. (1960/1980). *Probabilistic models for intelligence and attainment tests* (Copenhagen, Danish Institute for Educational Research). Expanded edition. Chicago, USA: The University of Chicago Press.

Reis, S.M., & Greene, M.J. (2002). *Using self-regulated learning to reverse underachievement in talented students*. Retrieved from http://www.gifted.uconn.edu/general/faculty/reis/Self-Regulated_Learning_Reverse_Underachievement.htm

Reis, S.M., & McCoach, D.B. (2000). The underachievement of gifted students: What do we know and where do we go? *Gifted Child Quarterly, 44*(3), 152-170. doi: 10.1177/001698610004400302

Ritchotte, J.A., Matthews, M.S., & Flowers, C.P. (2014). The validity of the achievement-orientation model for gifted middle school students an exploratory study. *Gifted Child Quarterly, 58*(3), 183-198. doi: 10.1177/0016986214534890

Snyder, K.E., & Linnenbrink-Garcia, L. (2013). A developmental, person-centered approach to exploring multiple motivational pathways in gifted underachievement. *Educational Psychologist, 48*(4), 209-228. doi: 10.1080/00461520.2013.835597

Tasmanian Qualification Authority (2006). *An introduction to Rasch Modelling and the TCE*. Retrieved from http://www.tqa.tas.gov.au/4DCGI/_WWW_doc/003675/RND01/Rasch_intro.pdf

Tasmanian Qualification Authority (2007). How the scaled awards are calculated and used to determine the tertiary entrance score. Retrieved from http://www.tqa.tas.gov.au/0477

Veas, A., Castejón, J.L., Gilar, R., & Miñano, P. (2015). Academic achievement in early adolescence: The influence of cognitive and non-cognitive variables. *The Journal of General Psychology, 142*(4), 273-294. doi: 10.1080/00221309.2015.1092940

Veas, A., Gilar, R., Miñano, P., & Castejón, J.L. (2016). Estimation of the Proportion of Underachieving students in Compulsory Secondary Education in Spain: an Application of the Rasch model. *Frontiers in Psychology, 7*, 303. doi: 10.3389/fpsyg.2016.00303

Valle, A., Regueiro, B., Estévez, I., Piñeiro, I., Rodríguez, S., & Freire, C. (2015). Homework involvement and motivation in elementary school students according to academic achievement and grade. *Europen Journal of Investigation in Health, Psychology and Education, 5*(3), 345-355.

Valle, A., Regueiro, B., Rodríguez, S., Piñeiro, I., Freire, C., Ferradás, M., & Suárez, N. (2015). Motivational profiles as a combination of self-efficacy expectations and academic goals in university students. *European Journal of Education and Psychology, 8*, 1-8.

Vlahovic-Stetic, V., Vidovic, V.V., & Arambasic, L. (1999). Motivational characteristics in mathematical achievement: A study of gifted high-achieving, gifted underachieving and non-gifted pupils. *High Ability Studies, 10*(1), 37-49. doi: 10.1080/1359813990100104

Wright, B.D., & Masters, G.N. (1982). *Rating scale analysis. Rasch measurement.* Chicago: MESA Press.

Wright, B.D., & Stone, M.H. (1979). *Best test design.* Chicago: MESA Press.

Wright, B.D., Linacre, J.M., Gustafson, J., & Martin-Lof, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), 370.

Yuste, C., Martínez, R., & Galve, J.L. (2005). *Batería de aptitudes diferenciales y generales. Badyg*. Madrid, Spain: CEPE.

Ziegler, A., & Stoeger, H. (2003). Identification of underachievement: An empirical study on the agreement among various diagnostic sources. *Gifted and Talented International, 18*(2), 87-94.

Ziegler, A., Ziegler, A., & Stoeger, H. (2012). Shortcomings of the IQ-based construct of underachievement. *Roeper Review, 34*(2), 123-132. doi: 10.1080/02783193.2012.660726

Received: March 23rd, 2016
Modifications Received: June 22nd, 2016
Accepted: June 24th, 2016