Azzopardi, L. M., Camilleri, L., Sammut, F., and Cefai, C. (2018). *Xjenza Online*, 6:16–24.

Xjenza Online - Science Journal of the Malta Chamber of Scientists
www.xjenza.org
DOI: 10.7423/XJENZA.2018.1.03

Research Article

Examining the Structural Validity of the Strengths and Difficulties Questionnaire (SDQ) in a Multilevel Framework

Lara Marie Azzopardi[1], **Liberato Camilleri**[*1], **Fiona Sammut**[1] **and Carmel Cefai**[2]
[1]*Department of Statistics and Operations Research, University of Malta, Msida, Malta*
[2]*Department of Psychology, University of Malta, Msida, Malta*

Abstract. The Strengths and Difficulties Questionnaire (SDQ), proposed by Goodman (1997), has been used by many researchers to measure the social, emotional and behaviour difficulties in children. The SDQ comprises four difficulty subscales measuring emotional, conduct, hyperactivity and peer problems. It also includes a fifth subscale measuring prosocial behaviour. A sample of 5200 Maltese students who were aged between 6 and 16 years was used to investigate the multilevel factor structure underlying the teachers' version of the SDQ. Statistical analysis in this study was conducted using Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA), Structural Equation Modelling (SEM) and Multilevel Structural Equation Modelling (MSEM). The study finds that a two-level three-factor model fits the data marginally better than a single-level three-factor model.

Keywords: Social emotional and behaviour difficulties (SEBD), Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA), Structural Equation Modelling (SEM), Multilevel Structural Equation Modelling (MSEM)

Abbreviations

Adjusted Goodness-of-Fit Index (AGFI), Comparative Fit Index (CFI), Goodness of Fit Index (GFI), Incremental Fit Index (IFI), Linear Structural Relations (LISREL), Kaiser Meyer Olkin (KMO), Modification Index (MI), Multiple Indicators and Multiple Causes (MIMIC), Multilevel Modelling (MLM), Normed Fit Index (NFI), Non-Normed Fit Index (NNFI), Preprocessor for LISREL (PRELIS), Principal Component Analysis (PCA), Relative Fit Index (RFI), Root Mean Square Error of Approximation (RMSEA), Strengths and Difficulties Questionnaire (SDQ), Weighted Least Squares Estimation (WLS), Parameter Estimate (Est.), Standard Error (S.E.).

1 Introduction

The SDQ, proposed by Goodman (1997), is a screening tool which is able to identify the prevalence of social, emotional and behaviour difficulties (SEBD) amongst children. The SDQ consists of five subscales that measure emotional, conduct, hyperactivity and peer problems together with prosocial behaviour. In turn, every subscale has five items all measured on a 3-point scale ranging from 0 to 2, where 0 corresponds to 'Not True', 1 to 'Somewhat True' and 2 corresponds to 'Certainly True'. Five of the items are negatively worded and require reverse-coding to generate the subscale scores, which range from 0 to 10. The total difficulty score, which excludes the prosocial subscale, ranges from 0 to 40. There exist three versions of the SDQ; one which is administered by the teacher, one by the parent and the other is self-administered by the student. These three SDQ versions have been translated into several languages, including Maltese. Cefai, Cooper and Camilleri (2008) validated the Maltese SDQ version through a process of forward and backward translations from English to Maltese. The reliability of the Maltese SDQ version was measured item by item where correlations ranged from 0.82 to 0.98.

2 Theoretical Framework

The main objective of this study is to analyse the factor structure underlying the rating scores provided to the 25 SDQ items by employing *Multilevel Structural Equation Modelling* (MSEM) which combines *Multilevel Modelling* (MLM) with *Structural Equation Modelling* (SEM).

Multilevel models (MLM) accommodate both fixed

**Correspondence to*: Liberato Camilleri (liberato.camilleri@um.edu.mt)

and random effects. Consequently, these models are appropriate to analyse clustered data that has a hierarchical nested structure. Multilevel models assume an error distribution at each level of nesting. With this modelling framework, it is possible to separate the observed variance within-clusters from the between-clusters components. On the other hand, the Structural Equation Modelling is a method which consists of three main analyses: Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA) and Structural Equation Modelling (SEM). The EFA accounts for patterns of correlations existing between the observable variables in terms of smaller number of latent variables. Once the EFA determines the factor structure, the CFA model is then fitted to the dataset to verify the pattern of the factor loadings, the number of the underlying factors and any covariances existing amongst the factors. Then, once the CFA confirms the latent structure, the SEM model is refitted to the data to analyse the relationships existing amongst the latent variables. SEM is made up of a measurement model and a structural model. A measurement model defines the latent constructs utilising various observable variables while a structural model assigns relationships between the latent variables.

In the presence of hierarchically nested data, Multilevel Structural Equation Modelling (MSEM) enables the researcher to fit either a CFA or SEM model at every nesting level of the hierarchical structure. This technique can be implemented by either starting with the MLM or the SEM analysis. In this paper, the statistical software Mplus (version 6) was used to analyse the data through MSEM. Mplus is statistical software which was developed by L. K. Múthen and Múthen (1998–2011). It is important to note that the MSEM methodology implemented by Mplus can only accommodate a two-level nesting structure. Consequently, in this analysis, a two-level model is fitted, where students (Level 1) are nested in classes (Level 2).

2.1 Model Specification of Two-level Factor Structure

Similarly to SEM, the multilevel structural equation model consists of the measurement model and the structural model. Additionally, the MSEM adjusts for the nested levels within the data by specifying a measurement model at each level and simultaneously enables the separation of the total variance into the within-group variance and the between-group variance. Rabe-Hesketh, Skrondal and Pickles (2004) defines a two-level SEM model as 'the specification of hierarchical conditional relationships. The response model specifies the distribution on the observed responses conditional on the latent variables and covariates (through a linear predictor and link function) and in the structural model the latent variables themselves may be regressed in other latent and observed covariates'.

The model specification of the two-level factor analysis follows the specification of Múthen's (1984, 1991, and 1994) work. This two-level confirmatory factor analysis decomposes the variability belonging to the indicators into the individual variability (within-level) and the between variability (between-level). This specification can be used with binary, ordinal categorical variables, censored, continuous responses or a combination of all. In this paper, this model assumes a two-level data structure existing with N statistical units (which in this dataset are students) which are in turn clustered in J groups (which in this dataset are classes). In this paper, it is assumed that the vector of the teachers' response can be decomposed into the sum of the within and between-class components.

Let $\mathbf{y}_{ij}$ be the vector of the teacher's SDQ responses about student i attending class j. By following Múthen's (1984) procedure, the model is constructed by defining an underlying normal distribution to the latent variable $\mathbf{y}^*_{pij}$ for the p^{th} observed variable $\mathbf{y}_{pij}$. Since the dataset consists of ordinal variables, this latent variable is defined by a set of thresholds. The vector of responses is illustrated as

$$\mathbf{y}^*_{ij} = \mathbf{y}^*_{B_g} + \mathbf{y}^*_{W_g}, \tag{1}$$

where $\mathbf{y}^*_{B_g}$ is the between-class contribution to the teachers' responses and $\mathbf{y}^*_{W_g}$ is the within-class contribution to the teachers' responses. Hence the total population covariance matrix ($\mathbf{\Sigma}_T$) can be decomposed into a between-class population covariance matrix ($\mathbf{\Sigma}_B$) and a within-class population covariance matrix ($\mathbf{\Sigma}_W$). Consequently,

$$\mathbf{\Sigma}_T = \mathbf{\Sigma}_W + \mathbf{\Sigma}_B. \tag{2}$$

Furthermore, it follows that

$$\mathbf{\Sigma}_W = \mathbf{\Lambda}_W \mathbf{\Phi}_W \mathbf{\Lambda}'_W + \mathbf{\Theta}_W, \tag{3}$$

and

$$\mathbf{\Sigma}_B = \mathbf{\Lambda}_B \mathbf{\Phi}_B \mathbf{\Lambda}'_B + \mathbf{\Theta}_B, \tag{4}$$

where $\mathbf{\Phi}_W$ and $\mathbf{\Phi}_B$ represent the covariance matrix for the within and the between-class factors. Moreover, $\mathbf{\Theta}_W$ and $\mathbf{\Theta}_B$ represent the covariance matrix for the diagonal matrices of the within and the between-class unique variance respectively.

2.2 The Two-level Measurement Model

Following Múthen's (1991) model, the model specification of the Level-1 (within-model) is defined as

$$\mathbf{y}_{ijk} = \boldsymbol{\alpha}_{jk} + \boldsymbol{\lambda}_{W_k} \boldsymbol{\eta}_{W_{ij}} + \boldsymbol{\varepsilon}_{W_{ijk}}, \tag{5}$$

where $\mathbf{y}_{ijk}$ represents the teacher's observed score on the indicator variable k given to student i attending class j;

$\boldsymbol{\alpha}_{jk}$ represents the intercept of indicator variable k in class j; $\boldsymbol{\lambda}_{W_k}$ represents the within-level factor loading $\boldsymbol{\lambda}_W$ of the indicator variable k; $\boldsymbol{\eta}_{W_{ij}}$ represents the score given by the teacher of student i attending class j on the within-level latent $\boldsymbol{\eta}_W$; $\boldsymbol{\varepsilon}_{W_{ijk}}$ represents the within-level error term , for the teacher of student i attending class j on the indicator variable k. Furthermore, the model specification of the Level-2 (between-model) is expressed as

$$\boldsymbol{\alpha}_{jk} = \boldsymbol{\nu}_k + \boldsymbol{\lambda}_{B_k}\boldsymbol{\eta}_{B_j} + \boldsymbol{\varepsilon}_{B_{jk}}, \tag{6}$$

where $\boldsymbol{\nu}_k$ represents the class-grand intercept of indicator variable k, which is the grand mean when the between-level latent variable is 0; $\boldsymbol{\lambda}_{B_k}$ refers to the between-level factor loadings of the indicator variable and $\boldsymbol{\eta}_{B_j}$ refers to the score of class j on the between-level latent variable $\boldsymbol{\eta}_B$; $\boldsymbol{\varepsilon}_{B_{jk}}$ refers to the between-level error term $\boldsymbol{\varepsilon}_B$ for class j on the indicator variable k. Consequently, one can deduce that

$$\mathbf{y}_{ijk} = \boldsymbol{\nu}_k + \boldsymbol{\lambda}_{B_k}\boldsymbol{\eta}_{B_j} + \boldsymbol{\varepsilon}_{B_{jk}} + \boldsymbol{\lambda}_{W_k}\boldsymbol{\eta}_{W_{ij}} + \boldsymbol{\varepsilon}_{W_{ijk}}. \tag{7}$$

The term $\boldsymbol{\alpha}_{jk}$ which is the class-specific item for the indicator variable k on the within-level is at the same time a dependent variable at the between-level (class-level). Consequently, this shows that the variability of class specific intercepts of an indicator variable k can be explained in the between-level by means of the latent variable $\boldsymbol{\eta}_{B_j}$. The non-explained variability in $\boldsymbol{\alpha}_{jk}$ is captured by the class-error term represented by $\boldsymbol{\varepsilon}_{B_{jk}}$.

2.3 The Two-level Latent Variable Model

The latent variable model of Level-1 can be represented by

$$\boldsymbol{\eta}_{W_{ij}} = \mathbf{B}_W\boldsymbol{\eta}_{W_{ij}} + \boldsymbol{\Gamma}_W\mathbf{X}_{W_{ij}} + \boldsymbol{\zeta}_{W_{ij}}. \tag{8}$$

$\mathbf{X}_{W_{ij}}$ consists of the vector of observed explanatory variables at the within-level; $\mathbf{B}_W$ consists of coefficient matrices between the latent and observed variables existing at the within-level; $\boldsymbol{\Gamma}_W$ consists of the vector of measurement intercepts at the within-level and $\boldsymbol{\zeta}_{W_{ij}}$ consists of the vector of disturbances at the within-level which are independent and randomly distributed with a zero mean and full variance-covariance matrix $\boldsymbol{\Theta}_W$ and $\boldsymbol{\Psi}_W$. Moreover, the latent variable model of Level-2 can be represented by

$$\boldsymbol{\eta}_{B_j} = \boldsymbol{\alpha}_B + \mathbf{B}_B\boldsymbol{\eta}_{B_j} + \boldsymbol{\Gamma}_B\mathbf{X}_{B_j} + \boldsymbol{\zeta}_{B_j}. \tag{9}$$

$\boldsymbol{\alpha}_B$ represents the class specific intercepts at the between levels; $\mathbf{X}_{B_j}$ consists of the vector of observed explanatory variables at the between-level; $\mathbf{B}_B$ consists of coefficient matrices between the latent and observed variables existing at the between-level; $\boldsymbol{\Gamma}_B$ consists of the vector of measurement intercepts at the between-level and $\boldsymbol{\zeta}_{B_j}$ consists of the vector of disturbances at the between-level which are independent and randomly distributed with mean 0 and full variance-covariance matrix $\boldsymbol{\Theta}_B$ and $\boldsymbol{\Psi}_B$.

3 Methodology

In this study, a random sample of 1326 teachers was selected from 110 schools, of which 66 were primary schools and 44 were secondary schools in order to analyse the social emotional and behaviour difficulties of 5200 students utilising the teacher SDQ version. Furthermore, the random sample, which was collected in 2005–2006, includes approximately 7% of the whole Maltese student population aged 6 to 16 years. In order to guarantee a representative sample, the students were stratified by gender, school-level, school-type and school region. The teachers were asked to assess the children they supervised by rating each of the 25 items of the Maltese SDQ teacher version. These items were measured on a 3-point scale ranging from 0 to 2 (not true, somewhat true and certainly true). The five items associated with emotional difficulties assessed anxiety, depression, fear and unhappiness. On the other hand, the five items associated with hyperactivity assessed restlessness, inattention, distraction, over-activity and inability to finish work. The five items associated with conduct problems assessed ill-temper and behaviour problems such as fighting, cheating, lying and stealing. The five items associated with peer problems assessed poor relations with peers, bullying and loneliness and the five items related to prosocial behaviour assessed good qualities like considerate, helpful, caring and kind to others. The scores of these five subscales were generated after reverse-coding the five negatively worded items.

4 Data Analysis and Results

The aim of this paper is to confirm whether or not a two-level three-factor model fits adequately the teachers' SDQ data collected in 2005–2006. Exploratory Factor Analysis, Confirmatory Factor Analysis, Structural Equation Modelling and Multilevel Structural Equation Modelling are utilised to identify the best model fit. EFA was carried out using SPSS software while CFA and SEM were carried out using LISREL (version 8.80) software. Furthermore, Mplus (version 6) statistical software was used to fit a two-level three-factor CFA model and a two-level three-factor SEM model for this SEBD dataset. The LISREL (version 8.80) software is not appropriate to fit a two-level CFA/SEM model because it accommodates continuous responses but not ordinally-scaled categorical responses.

4.1 Internal Consistency

Cronbach Alpha (Cronbach, 1951) was utilised to assess the internal consistency of the items within every subscale. The items in the Conduct, Hyperactivity, Emotional and Prosocial subscales have satisfactory internal consistency and their Cronbach Alpha exceeded the 0.7 threshold values. The Peer subscale had a weak internal consistency, since its Cronbach Alpha just exceeded the 0.5 threshold. The item *Child gets on better with adults than children of same age* was weakly related to other items in this subscale.

4.2 Exploratory Factor Analysis

Exploratory Factor Analysis was used to assess the factorial validity of the whole SEBD data and to identify the number of latent dimensions underlying this dataset. The Kaiser Meyer Olkin (KMO) value, which gives an indication of the relative compactness of the correlations, was found to be equal to 0.898, which exceeds the 0.5 threshold value. Moreover, the Bartlett's test of sphericity, which tests whether the correlation matrix is significantly different from the identity matrix, yielded a p-value less than the 0.05 level of significance. Both results indicate a latent structure within this SEBD data and that EFA is essential to reveal the latent factor structure.

Table 1 displays the factor loadings of this three-factor model. Stevens (2002) suggested a threshold value of 0.4 for these factor loadings when the sample size exceeds 150 observations and the number of variables exceeds 10. Factor 1, which represents the Externalisation dimension, comprises nine of the items in the Hyperactivity and Conduct subscales, including *Tempers, Obedient, Fights, Lies, Restless, Distractible, Fidgety, Reflective* and *Persistent.* Moreover, Factor 2, which comprises all the items in Prosocial subscale, includes *Considerate, Shares, Caring, Kind to kids* and *Helps out.* Furthermore, Factor 3, which represents the Internalisation dimension, comprises six of the items in the Emotion and Peer subscales, including *Worries, Unhappy, Clingy, Fears, Solitary* and *Bullied.*

4.3 Confirmatory Factor Analysis (CFA)

A three-factor CFA model was then fitted to the whole SEBD sample using the Weighted least squares (WLS) estimation technique. This is the appropriate estimation technique when analysing ordinal categorical responses (rating scores). The fitted model defines the relationships amongst the Externalisation, Prosocial and Internalisation dimension while relaxing some of the assumptions posed in EFA. Once a model was specified, the t-rule was used to assess whether the model is identified. Since the t-value for the model fit was found to be 45, which is less than the $0.5q(q+1) = 210$ criterion, then the three-factor CFA model has model identification. The resulting parameter estimates of lambda-x, phi-paths and theta-deltas were all found to be significant since the corresponding z-scores exceed 1.96 for all observed variables.

Table 1: The factor loadings of 3 factors using Varimax Rotation.

Variable Description	Externalisation Factor	Prosocial Factor	Internalisation Factor
Tempers	**0.54**	0.20	0.08
Obedient	**0.56**	0.27	0.10
Fights	**0.62**	0.21	−0.01
Lies	**0.55**	0.27	0.09
Steals	0.22	0.12	0.11
Somatic	0.21	0.03	0.38
Worries	−0.04	−0.05	**0.62**
Unhappy	0.16	0.08	**0.60**
Clingy	0.12	0.06	**0.60**
Fears	−0.03	0.04	**0.67**
Restless	**0.77**	−0.04	−0.07
Fidgety	**0.80**	0.01	0.01
Distractible	**0.60**	0.20	0.27
Reflective	**0.54**	0.35	0.15
Persistent	**0.48**	0.31	0.27
Solitary	−0.11	0.25	**0.45**
Good Friend	0.05	0.25	0.22
Popular	−0.03	0.39	0.28
Bullied	0.09	0.12	**0.42**
Best with adults	0.12	−0.02	0.20
Considerate	−0.34	**−0.67**	−0.01
Shares	−0.20	**−0.64**	−0.06
Caring	−0.18	**−0.69**	−0.03
Kind to kids	−0.24	**−0.61**	−0.02
Helps out	−0.21	**−0.66**	−0.05

Fig. 1 displays the path diagram and the corresponding WLS estimates of the three-factor CFA model. The path diagram shows the relationships between the three dimensions (Externalisation Internalisation and Prosocial factors) and their relationships with the twenty observed items. The Externalisation factor respectively explained 84%, 77%, 76%, 74% and 71% of the variances of the items *Distractible, Fight, Persistent, Lies* and *Tempers.* The Internalisation factor explains 91% of the variance of the item *Unhappy* and the Prosocial factor explains 79% of the variance of the item *Considerate.* The majority of the standardized factor loadings exceed 0.7, indicating that the latent factors strongly affect 18 of the observed variables and moderately affect the remaining 2 items: *Solitary* and *Bullied.* Furthermore, the CFI (0.93), GFI (0.98), AGFI (0.97), NFI (0.92), NNFI (0.92), IFI (0.93) and RFI (0.91) all exceed their threshold values indicating a well-fitted model. Moreover, the Hoelter's Critical N (393.5) exceeds the 200 cut-point and the RMSEA value (0.06) is less than the 0.07 threshold value suggested by Steiger (2007). All these fit indices satisfy their threshold criteria, which indicate that this three-factor CFA model (Model 1) fits the data well.

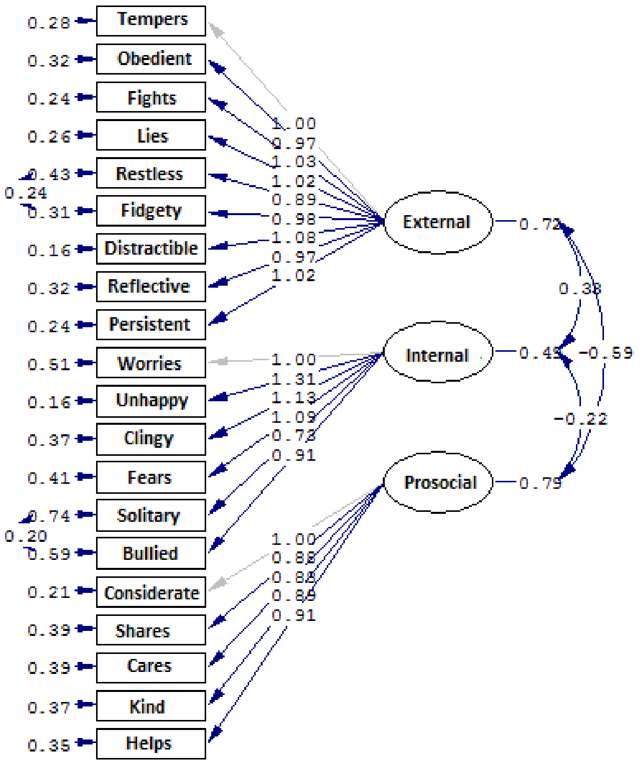

Figure 1: Path diagram of the three-factor CFA Model (Model 1).

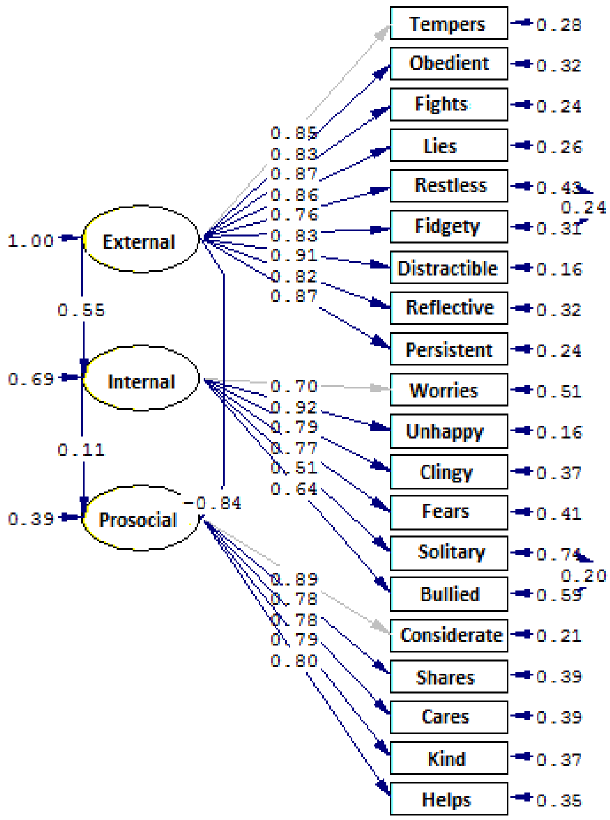

Figure 2: Path diagram of the three-factor SEM Model (Model 2).

4.4 Structural Equation Modelling (SEM)

A three-factor structural equation model was also fitted on this SEBD dataset using LISREL (version 8.8) software to investigate the relationships existing between the latent variables. Essentially this involves regressing latent variables on one another.

Fig. 2 displays the path diagram of this three-factor SEM model, which shows the relationships between the Externalisation, Internalisation and Prosocial factors and their relationships with the 20 observed items. Once the model was specified, the t-rule was used to check that this three-factor SEM fitted has model identification. The model parameters were estimated using a weighted least squares estimation (WLS) technique. The corresponding factor loadings, phi-paths and theta-deltas estimates are all significant since their standard errors are less than half the value of the parameter estimates. The CFI (0.92), GFI (0.98), AGFI (0.97), NFI (0.92), NNFI (0.91), IFI (0.92) and RFI (0.91) all exceed their threshold values by a small margin indicating a plausible fit.

This model can be extended further by accommodating the nesting structure of the data, where individuals are nested in classes.

4.5 A CFA Model in a Multilevel Framework

By fitting a two-level three-factor CFA model on this dataset, one can now estimate the between-level and the within-level loadings. This model explains the relationship existing at the class-level (Level-2) and at the individual-level (Level-1). In contrast the CFA model shown in Fig. 1 ignores this variation amongst the units. From this analysis, it was found that the within-level correlations range from −0.567 to 0.767, whilst the between-level correlations range from −0.572 to 0.830.

The three-factor model displayed in Fig. 1 was fitted at both Level-1 (student level) and at Level-2 (class-level). In this study, a WLSMV estimation technique was utilised following Múthen's (1984) recommendation. The WLSMV estimation technique is the most suitable estimation technique utilised for ordinal data. Furthermore, the WLSMV estimation technique for this analysis was carried out using the delta parameterisation. The delta parameterisation sets the measurement residuals equal to a value of 1. As pointed out by Newsom (2014), this parameterisation can be considered to be a variant of the probit model.

The number of clusters existing in this analysis was 1321 and the quasi-average cluster size was found to

Table 2: The estimated Intra-class Correlation Coefficient (ICCs).

Variable Description	Intra-Class Correlation
Tempers	0.237
Obedient	0.214
Fights	0.189
Lies	0.204
Restless	0.153
Fidgety	0.138
Distractible	0.180
Reflective	0.195
Persistent	0.208
Worries	0.230
Unhappy	0.221
Clingy	0.249
Fears	0.228
Solitary	0.131
Bullied	0.220
Considerate	0.248
Shares	0.333
Cares	0.329
Kind	0.380
Helps	0.280

be equal to 3.936. Table 2 displays the intra-class correlation coefficients of these twenty observed variables. This table shows that all the intra-class correlations are greater than 0.10 illustrating that the effects of classes are influencing the teachers' rating scores. As pointed out by Gajewski, Boyle, Miller, Oberhelman and Dunton (2010), this further shows that the multilevel CFA is appropriate for particular dataset. Furthermore this result also indicates that the effects of class are strongly influencing the SEBD rating scores provided by the teachers.

The fitted two-level three-factor CFA model is considered to be identified since the same factor structure holds in both the within and the between-level. Table 3 displays the unstandardised factor loadings at the within-level, where all loadings are significant. Moreover, the mean standardised loading of the observed variables at the within-level was found to be equal to 0.7283.

Table 4 shows the unstandardised factor loadings at the between-level, where nearly all loadings are significant. Moreover, the mean standardised loading of the observed variables at the between-level was found to be equal to 0.7736. The fit indices showed that this model fits the data adequately well. The CFI (0.839) and TLI (0.815) all exceed their threshold values by a small margin indicating a plausible fit. Furthermore, the Hoelter's Critical N (325.96) exceeds the 200 cut-point and the RMSEA value was (0.059), which is less than the 0.07 threshold value suggested by Steiger (2007). Additionally, the SRMR value of the within-level was (0.108), whilst the SRMR value of the between-level was (0.107). This indicates that this model fits slightly better the data at Level-1 (Individual-level) than at Level-2 (Class-level).

Table 3: Factor loadings at Level-1 of the CFA model. Here, Est. represents the Parameter Estimate and S.E. is the Standard Error.

Within-level	**Est.**	**S.E.**	**Est./S.E.**	***p*-value**
Externalisation by				
C1			Aliased	
C2	1.298	0.066	19.571	0.000
C3	1.585	0.089	17.763	0.000
C4	1.358	0.076	17.948	0.000
C5	0.878	0.045	19.601	0.000
C6	1.089	0.053	20.579	0.000
C7	1.524	0.074	20.576	0.000
C8	1.415	0.072	19.651	0.000
C9	1.466	0.074	19.817	0.000
Internalisation by				
E1			Aliased	
E2	1.868	0.144	12.972	0.000
E3	1.466	0.092	16.019	0.000
E4	1.693	0.109	15.468	0.000
E5	0.892	0.062	14.481	0.000
E6	1.073	0.079	13.586	0.000
Prosocial by				
P1			Aliased	
P2	0.533	0.033	16.263	0.000
P3	0.571	0.034	16.960	0.000
P4	0.593	0.038	15.488	0.000
P5	0.614	0.037	16.622	0.000
Externalisation with				
Internalisation	0.083	0.015	5.626	0.000
Prosocial	−1.128	0.079	−14.279	0.000
Internalisation with				
Prosocial	−0.21	0.032	−6.549	0.000
C5 WITH				
C6	0.767	0.011	69.774	0.000
E5 WITH				
E6	0.222	0.035	6.343	0.000
Variances				
Externalisation	0.839	0.068	12.277	0.000
Internalisation	0.508	0.044	11.565	0.000
Prosocial	3.286	0.318	10.347	0.000

The mean standardised loading of the observed vari-

Table 4: Factor loadings at Level-2 of the CFA model.

Between-level	Est.	S.E.	Est./S.E.	p-value
Externalisation by				
C1			Aliased	
C2	1.072	0.105	10.243	0.000
C3	1.133	0.112	10.127	0.000
C4	1.129	0.108	10.462	0.000
C5	0.627	0.066	9.492	0.000
C6	0.781	0.074	10.495	0.000
C7	1.179	0.105	11.278	0.000
C8	1.018	0.103	9.855	0.000
C9	1.068	0.107	10.014	0.000
Internalisation by				
E1			Aliased	
E2	1.449	0.150	9.682	0.000
E3	1.568	0.158	9.921	0.000
E4	1.505	0.141	10.640	0.000
E5	0.651	0.092	7.039	0.000
E6	1.031	0.145	7.111	0.000
Prosocial by				
P1			Aliased	
P2	0.788	0.060	13.140	0.000
P3	0.774	0.056	13.760	0.000
P4	0.866	0.067	12.960	0.000
P5	0.670	0.050	13.340	0.000
Externalisation with				
Internalisation	0.236	0.031	7.605	0.000
Prosocial	−0.361	0.052	−6.947	0.000
Internalisation with				
Prosocial	−0.191	0.033	−5.788	0.000
C5 WITH				
C6	0.100	0.027	3.674	0.000
E5 WITH				
E6	−0.013	0.028	−0.461	0.645
Variances				
Externalisation	0.315	0.051	6.119	0.000
Internalisation	0.207	0.035	5.838	0.000
Prosocial	1.274	0.162	7.842	0.000

ables at the between-level was found to be higher than the mean standardised loading of the observed variables observed at the within-level. The reason behind this fact is that the between-level is based on the means. Consequently, these means are more reliable when compared to the raw scores and in this case a lot of measurement error was eliminated. It was also noticed that generally, the obtained standardised factor loadings are substantially higher at the class-level than the individual-level. Tables 5 and 6 show the obtained multiple correlations (R^2) values obtained at the within-level and at the between-level respectively. These values suggest the strength of every observed variable in measuring the corresponding factor at each level. Similarly to the standardised factor loadings, the multiple correlations values are in general higher at the class-level than the individual-level. As shown in Tables 3 and 4, the strongest item in measuring the Externalisation Factor is the item *Fights* at the within-level and the item *Distractible* at the between-level. Moreover, the strongest item in measuring the Internalisation Factor is the item *Unhappy* at the within-level and the item *Clingy* at the between-level. Furthermore, the strongest item in measuring the Prosocial Factor in both the within and between-level is the item *Considerate.*

Table 5: The obtained multiple correlations values of the within-level.

	Within-Level			
Variable	Estimate	S.E.	Est./S.E.	p-value
Tempers	0.456	0.020	22.575	0.000
Obedient	0.586	0.018	31.912	0.000
Fights	0.678	0.021	33.088	0.000
Lies	0.607	0.021	29.220	0.000
Restless	0.393	0.017	22.750	0.000
Fidgety	0.499	0.017	30.118	0.000
Distractible	0.661	0.014	45.811	0.000
Reflective	0.627	0.015	42.847	0.000
Persistent	0.643	0.015	42.901	0.000
Worries	0.337	0.019	17.440	0.000
Unhappy	0.639	0.030	21.261	0.000
Clingy	0.522	0.023	23.179	0.000
Fears	0.593	0.024	24.322	0.000
Solitary	0.288	0.024	12.011	0.000
Bullied	0.369	0.028	13.303	0.000
Considerate	0.767	0.017	44.348	0.000
Shares	0.483	0.019	25.314	0.000
Cares	0.517	0.018	28.249	0.000
Kind	0.536	0.021	25.179	0.000
Helps	0.554	0.018	30.239	0.000

If ψ_W and ψ_B are the within-level and the between-level variances then the intra-class correlations for the latent externalisation, internalisation and prosocial factors are

$$ICC_E = \rho_E = \frac{\psi_B}{\psi_B + \psi_W} = \frac{0.315}{0.315 + 0.839} = 0.273, \tag{10}$$

$$ICC_I = \rho_I = \frac{\psi_B}{\psi_B + \psi_W} = \frac{0.207}{0.207 + 0.508} = 0.290, \tag{11}$$

Table 6: The multiple correlations values of the between-level.

	Between-Level			
Variable	Estimate	S.E.	Est./S.E.	p-value
Tempers	0.550	0.056	9.844	0.000
Obedient	0.549	0.054	10.204	0.000
Fights	0.558	0.068	8.151	0.000
Lies	0.616	0.065	9.534	0.000
Restless	0.415	0.067	6.198	0.000
Fidgety	0.602	0.071	8.422	0.000
Distractible	0.675	0.048	14.057	0.000
Reflective	0.503	0.049	10.332	0.000
Persistent	0.488	0.045	10.763	0.000
Worries	0.459	0.065	7.003	0.000
Unhappy	0.553	0.069	8.057	0.000
Clingy	0.735	0.071	10.300	0.000
Fears	0.648	0.075	8.643	0.000
Solitary	0.415	0.096	4.313	0.000
Bullied	0.492	0.087	5.665	0.000
Considerate	0.900	0.054	16.652	0.000
Shares	0.820	0.057	14.268	0.000
Cares	0.750	0.049	15.257	0.000
Kind	0.723	0.055	13.146	0.000
Helps	0.656	0.049	13.267	0.000

$$ICC_P = \rho_P = \frac{\psi_B}{\psi_B + \psi_W} = \frac{1.274}{1.274 + 3.286} = 0.279. \tag{12}$$

4.6 A SEM Model in a Multilevel Framework

A two-level three-factor SEM model was also fitted on this SEBD dataset by using the software Mplus (version 6). A two-level three-factor SEM model illustrates the relationships existing amongst the three factors: Externalisation, Internalisation and Prosocial Factors. Following, the analysis of fitting a SEM model, it was intuitively observed that in this two-level model, the Internalisation Factor depends on the Externalisation Factor, the Prosocial Factor depends on the Externalisation Factor and the Prosocial Factor depends on the Internalisation Factor. This fitted two-level three-factor SEM model is considered to be identified since the same factor structure holds in both the within and the between-level. Tables 7 and 8 display the unstandardised factor loadings, standard errors, Wald statistics and p-values at the within-level and the between-level of the three-factor SEM model.

All the obtained unstandardised and standardised factor loadings of the two-level SEM model at the within-level are considered to be significant since their standard errors are less than the half the value of the loadings. Moreover, their Wald test statistics are greater than $|1.96|$ and their corresponding p-values are less than 0.05. On the other hand, nearly all the obtained unstandardised and standardised factor loadings at the between-level are significant since their standard errors are less than the half the value of the loadings with the exception of the parameter estimate existing amongst the observed variables *Bullied* and *Solitary*, the residual variance of the item *Considerate* and the residual variance of the *Externalisation Factor*.

Table 7: Factor loadings at Level-1 of the CFA model.

Within-level	**Est.**	**S.E.**	**Est./S.E.**	***p*-value**
Externalisation by				
C1			Aliased	
C2	1.298	0.066	19.570	0.000
C3	1.585	0.089	17.760	0.000
C4	1.358	0.076	17.950	0.000
C5	0.878	0.045	19.600	0.000
C6	1.089	0.053	20.580	0.000
C7	1.524	0.074	20.580	0.000
C8	1.415	0.072	19.650	0.000
C9	1.466	0.074	19.820	0.000
Internalisation by				
E1			Aliased	
E2	1.868	0.144	12.970	0.000
E3	1.466	0.092	16.020	0.000
E4	1.693	0.109	15.470	0.000
E5	0.892	0.062	14.480	0.000
E6	1.073	0.079	13.590	0.000
Prosocial by				
P1			Aliased	
P2	0.533	0.033	16.260	0.000
P3	0.571	0.034	16.960	0.000
P4	0.593	0.038	15.490	0.000
P5	0.614	0.037	16.620	0.000
Externalisation with				
Internalisation	0.021	0.024	0.886	0.376
Prosocial	−0.342	0.023	−15.090	0.000
Internalisation with				
Prosocial	−0.206	0.010	−6.305	0.000
C5 WITH				
C6	0.767	0.011	69.770	0.000
E5 WITH				
E6	0.222	0.035	6.343	0.000
Variances				
Externalisation	0.452	0.037	12.140	0.000
Internalisation	0.495	0.043	11.440	0.000
Prosocial	3.286	0.318	10.350	0.000

Table 8: Factor loadings at Level-2 of the SEM model.

Between-level	Est.	S.E.	Est./S.E.	p-value
Externalisation by				
C1			Aliased	
C2	1.072	0.105	10.240	0.000
C3	1.133	0.112	10.130	0.000
C4	1.129	0.108	10.460	0.000
C5	0.627	0.066	9.492	0.000
C6	0.781	0.074	10.490	0.000
C7	1.179	0.105	11.280	0.000
C8	1.018	0.103	9.855	0.000
C9	1.068	0.107	10.010	0.000
Internalisation by				
E1			Aliased	
E2	1.449	0.150	9.682	0.000
E3	1.568	0.158	9.921	0.000
E4	1.505	0.141	10.640	0.000
E5	0.651	0.092	7.039	0.000
E6	1.031	0.145	7.111	0.000
Prosocial by				
P1			Aliased	
P2	0.788	0.060	13.140	0.000
P3	0.774	0.056	13.760	0.000
P4	0.866	0.067	12.960	0.000
P5	0.670	0.050	13.340	0.000
Externalisation with				
Internalisation	1.021	0.137	7.443	0.000
Prosocial	−0.131	0.035	−3.744	0.000
Internalisation with				
Prosocial	−0.150	0.025	−5.906	0.000
C5 WITH				
C6	0.100	0.027	3.674	0.000
E5 WITH				
E6	−0.013	0.028	−0.461	0.645
Variances				
Externalisation	0.026	0.029	0.890	0.374
Internalisation	0.179	0.032	5.576	0.000
Prosocial	1.274	0.162	7.842	0.000

5 Conclusion

Table 9 shows the fit indices obtained for the single-level three-factor SEM models compared to the fit indices obtained for the two-level three-factor SEM model. Although the two-level SEM model reduced the CFI and RMSEA indices it did not yield the significant improvement that was expected considering the complexity of model. The reason is that the level-1 variance is considerably larger than the level-2 variance. This two-level SEM model would have been more appropriate if the level-2 variance explained a larger portion of the total variance.

Table 9: Fit indices of the 1-level and 2-level SEM models.

Fit Index	Single-level SEM	Two-level SEM
CFI	0.920	0.839
RMSEA	0.061	0.059

References

Cefai, C., Cooper, P. & Camilleri, L. (2008). *Engagement Time A National study of students with social, emotional and behaviour difficulties in Maltese schools.* San Gwann, Malta: European Centre for Educational Resilience and Socio-Emotional Health, University of Malta.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of testes. *Psychometrika*, *16*(3), 297–334.

Gajewski, B., Boyle, D., Miller, P., Oberhelman, F. & Dunton, N. (2010). A Multilevel Confirmatory Factor Analysis Framework. Psychological Methods. *19*(1), 72–91.

Goodman, R. (1997). Strengths and Difficulties Questionnaire: Research Note. *J. Child Psychol. Psychiatry*, *38*, 581–586.

Múthen, B. O. (1984). A general structual equation model with dichotomous, ordered categorical and continuous latent variables indicators. *Psychometrika*, *49*, 115–132.

Múthen, B. O. (1991). Multilevel factor analysis of class and student acheivement components. *J. Educ. Meas.* *28*(4), 254–338.

Múthen, B. O. (1994). Multilevel covariance structure analysis. *Sociol. Methodol.* *22*(3), 376–398.

Múthen, L. K. & Múthen, B. O. (1998–2011). Mplus (Version 6) [Computer Program]. Los Angeles, CA: Muthen & Muthen.

Newsom, J. T. (2014). *Longitudinal Structural Equation Modeling: A Comprehensive Introduction.* New York: Routledge.

Rabe-Hesketh, S., Skrondal, A. & Pickles, A. (2004). Generalized multilevel structural equation modeling. *Psychometrika*, *69*(2), 167–190.

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personal. Individ. Differ.* *42*(5), 893–898.

Stevens, J. P. (2002). *Applied Multivariate Statistics for the Social Sciences* (4th Edition). Lawrence Erlbaum Associates.