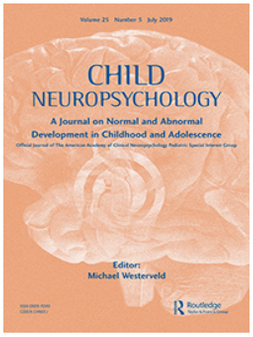

Routledge
Taylor & Francis Group

Child Neuropsychology

A Journal on Normal and Abnormal Development in Childhood and Adolescence

ISSN: 0929-7049 (Print) 1744-4136 (Online) Journal homepage: https://www.tandfonline.com/loi/ncny20

Dimensional structure and measurement invariance of the BRIEF-2 across gender in a socially vulnerable sample of primary school-aged children

Angie Jiménez & Beatriz Lucas-Molina

To cite this article: Angie Jiménez & Beatriz Lucas-Molina (2019) Dimensional structure and measurement invariance of the BRIEF-2 across gender in a socially vulnerable sample of primary school-aged children, Child Neuropsychology, 25:5, 636-647, DOI: 10.1080/09297049.2018.1512962

To link to this article: https://doi.org/10.1080/09297049.2018.1512962

CHILD NEUROPSYCHOLOGY
2019, VOL. 25, NO. 5, 636–647
https://doi.org/10.1080/09297049.2018.1512962

Routledge
Taylor & Francis Group

Dimensional structure and measurement invariance of the BRIEF-2 across gender in a socially vulnerable sample of primary school-aged children

Angie Jiménez and Beatriz Lucas-Molina

Department of Developmental and Educational Psychology, University of Valencia, Valencia, Spain

ABSTRACT

The present study aims to analyze, on the one hand, the dimensional structure of the BRIEF-2 and, on the other, its measurement invariance across gender in a sample of primary school-aged children. Participants were 168 students (51.8% girls) in conditions of social risk and in foster homes, with ages ranging from 6 to 15 years old (M = 10.08; *SD* = 2.09). The children's executive functioning level was obtained from their parents or caregivers. The analysis of the internal structure using confirmatory factor analysis reveals that the model with three indexes (Behavioral, Emotional and Cognitive) and nine scales (Inhibit, Shift, Self-Monitor, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials and Task-Monitor) showed a good fit to the sample. Likewise, the results support the measurement invariance of the BRIEF-2 across genders. The coefficient alpha values for the index scores ranged from .71 to .91, with coefficients for the individual scales ranging from .66 to .88. The findings support the factorial validity of the BRIEF-2 scores, suggesting that it is an adequate instrument to evaluate executive functioning reported by parents and caregivers. Future studies must continue to examine the structure and invariance of the BRIEF-2 in different populations and cultures.

ARTICLE HISTORY
Received 20 April 2018
Accepted 12 August 2018

KEYWORDS
BRIEF-2; executive functions; measurement invariance; gender; socially vulnerable children

Executive functioning is the capacity to manage and optimize cognitive and emotional resources through continuous processes of planning, systematization, regulation, control, and monitoring, in order to resolve complex or novel situations and achieve the target behavior. According to Gioia, Isquith, Kenworthy, and Barton (2002), this capacity involves the use of related behavioral, emotional, and metacognitive skills.

From a multidimensional perspective, the executive functions are considered higher-order skills in neuropsychological development (Stelzer, Mazzoni, & Cervigni, 2013; Zelazo and Carlson, 2012) that make it possible to act in a conscious, willful, intentional, and flexible way to situations requiring decision-making. Therefore, they represent key factors for people's adequate coping in daily life and their adaptation to new challenges or objectives (Abundis, 2014; Barkley, 2012; Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008). Many researchers highlight that the term "executive function" is

CONTACT Beatriz Lucas-Molina beatriz.lucas@uv.es Department of Developmental and Educational Psychology, University of Valencia, Avenida Blasco Ibáñez 21, Valencia 46010, Spain

an umbrella construct (Dodzik, 2017) that groups key or primary skills such as inhibition, working memory, cognitive flexibility, planning, self-regulation, initiation, problem-solving, and the continuous use of feedback and monitoring (Anderson, 2002; Lyons Usher, Leon, Stanford, Holmbeck, & Bryant, 2016; McDermott, Westerlund, Zeanah, Nelson & Fox, 2012; Obradović, 2016, Zelazo et al., 2003).

In the study of these neuropsychological skills, the BRIEF (The Behavior Rating Inventory of Executive Function, Gioia, Isquith, Guy, & Kenworthy, 2000) is a relevant, useful, and ecological assessment test, both in its original version and in the second revision known as the BRIEF-2 (Gioia, Isquith, Guy, & Kenworthy, 2015). These scales make it possible to explore the behavioral, emotional, and cognitive dimensions of executive functioning in the practical and natural life of children and adolescents, based on reports by their parents, caregivers, or teachers (Fournet et al., 2015; Gioia et al., 2000, 2015; Gioia, Isquith, Retzlaff, & Espy, 2002). This test is widely used in the clinical setting and with the general pediatric population (Jacobson, Pritchard, Koriakin, Jones, & Mahone, 2016).

The inventory has various versions to evaluate executive functioning depending on the ages of the children to be assessed (under or over the age of 5). In the case at hand, we will explore the parent version used to evaluate children and adolescents from 5 to 18 years old. In its original version, the BRIEF contained 86 items (Gioia et al., 2000), grouped in eight scales and two indexes: the Behavioral Regulation Index (BRI), which includes the Emotional Control, Shift, and Inhibit scales, and the Metacognition Index (MI), which includes the Plan/Organize, Working Memory, Initiate, Organization of Materials, and Monitor scales. These two indexes are grouped in a Global Executive Composite (GEC).

It is important to point out that the BRIEF has been used in various studies, including some that analyze the type of internal structure and the factorial grouping that best fits this structure (e.g., Donders, DenBraber, & Vos, 2010; Egeland and Fallmyr, 2010; Fournet et al., 2015; Gioia, Isquith, Retzlaff et al., 2002; Lyons Usher et al., 2016; Memisevic, 2015; Peters, Algina, Smith, & Daunic, 2012; Slick, Lautzenhiser, Sherman & Eyrl, 2006). In this regard, the original structure with two factors and eight scales has been confirmed in different studies with clinical samples (e.g., Donders et al., 2010; Lyons Usher et al., 2016; Memisevic, 2015; Slick et al., 2006). However, these studies have also shown some differences in the way and strength with which the different scales saturated in the BRI and MI. For example, the Monitoring subscale was found to present the same factorial load in both indexes (Slick et al., 2006). This result was also found when testing an oblique two-factor model (Lyons Usher et al., 2016), which would mean that this scale can have a component of task monitoring and another of behavior monitoring.

In 2015, the BRIEF-2 (Gioia et al., 2015) was published, with important differences from the original version published in the year 2000, such as the reduction in the number of items (from 86 to 63, in an attempt to design a shorter version) and the validation of a structure with three indexes and nine scales, instead of two indexes and eight scales. Thus, the Shift and Emotional Control scales are grouped in a new Emotional Regulation Index (ERI), which explores the capacity to regulate, monitor, and have flexible emotional responses to problems that arise. The Inhibit and Self-Monitor scales are grouped in the Behavior Regulation Index (BRI), which examines the child's capacity to self-regulate and

control his/her behavior. Finally, the Initiate, Working Memory, Plan/Organize, Task Monitoring, and Organization of Materials scales make up the Cognitive Regulation Index (CRI), which studies the capacity to manage and regulate the cognitive processes in order to achieve goals or deal with problems faced (Dodzik, 2017; Gioia et al., 2015). In this version, the monitoring scale has been divided into a personal monitoring dimension and another task monitoring dimension, as indicated in previous studies (e.g., Lyons Usher et al., 2016; Slick et al., 2006).

It is important to point out that, although the new structure of the BRIEF-2 is recent, evidence had already been obtained for the grouping of three factors and nine scales in the original BRIEF for both the parent (Egeland and Fallmyr, 2010; Fournet et al., 2015; Gioia, Isquith, Retzlaff et al., 2002) and teacher versions (Egeland and Fallmyr, 2010; Peters et al., 2012).

For example, Gioia, Isquith, Retzlaff et al. (2002) tested the fit of four models of the BRIEF structure in the parents' version with a clinical population. They examined a one-dimensional model, a two-factor model with eight scales (the original structure of the BRIEF), another model with three factors and nine scales (the later structure of the BRIEF-2), and a final model with four factors in which the Emotional control scale was separate from the BRI. They found that the three-factor structure with nine scales presented the best internal fit. It should be pointed out that all the models, including the one with three factors, showed root mean square error of approximation (RMSEA) values above .11, exceeding the recommended value for model acceptance (.08, Hu and Bentler, 1995). Later, Egeland and Fallmyr (2010) also explored these models, except the four-factor model, in the parents' and teachers' versions. They obtained results similar to those found by Gioia, Isquith, Retzlaff et al. (2002), with the three-factor model with nine scales presenting the best fit, followed by the model with two factors and eight scales.

In addition, Fournet et al. (2015) also explored different models (one-dimensional with eight and nine scales, two factors with eight scales, and three factors with nine scales), and they also found that the two-factor structure with eight scales and the three-factor structure with nine scales in the original version for parents showed a good fit. However, for the teachers' version, good fit levels were not obtained. These researchers also confirmed the structural invariance of the model with three factors and nine scales across gender. Thus, they found that girls presented greater executive efficiency on all the scales, except Organization of Materials and Monitoring.

Few studies have validated the three-factor structure with nine scales of the parents' and teachers' versions of the recent BRIEF-2. In fact, to our knowledge, only Maldonado, Fournier, Martínez, González Espejo-Saavedra and Santamaría (2017) carried out a Spanish adaptation and validation of the BRIEF-2. These authors found a moderate fit of the model with three factors and nine scales. The RMSEA index showed values of .09 and .16, again above the values considered acceptable (.08, Hu and Bentler, 1995), both in a clinical sample ($n > 200$) and in the standard sample ($n > 1500$) used for the adaptation of the two versions of the test.

As can be observed, different models of the factorial grouping of the BRIEF in its two versions have been explored. However, the model of the original version of the BRIEF (two factors and eight scales) and the model the BRIEF-2 presents now (three factors and nine scales) have received the most empirical support so far. The fact that not all

the studies have been successful in confirming the BRIEF structures continues to be a topic of debate in the research community. Likewise, the new version of the test has not been widely addressed, and so there is a need to continue to research in order to better understand the ecological behavior of the executive functions measured through tests as widely employed tests as the BRIEF.

Furthermore, gender has been one of the most debated topics in the study of the executive functions in childhood. Hence, better executive functioning has been observed in girls (Fournet et al., 2015; Huizinga and Smidts, 2011; Maldonado, 2016), especially on the scales of Inhibit, Plan/Organize, Initiate, Organization of Materials, and Monitor. However, to the best of our knowledge, only one study (Fournet et al., 2015) has examined the invariance of the original version of the BRIEF across gender, as a prior step in the analysis of possible differences in the executive functioning of boys and girls.

Various studies have pointed to the impact of conditions of risk and social vulnerability and negative experiences on executive functioning. Boys and girls who have experienced situations of abuse, poverty, abandonment, and traumas tend to demonstrate less development of the executive functions, according to a variety of studies (e.g., Berthelsen, Hayes, White, & Williams, 2017; Blair and Raver, 2012; Masten et al., 2012; Obradović, 2016; Raver, Blair, & Willoughby, 2013; Roos, Kim, Schnabler, & Fisher, 2016). Therefore, the importance of developing valid instruments to evaluate the executive functions in childhood becomes even greater in the case of populations in situations of poverty and social vulnerability.

Within this research framework, the present study aims to analyze, on the one hand, the dimensional structure of the BRIEF-2 and, on the other, its measurement invariance across gender in a sample of primary school-aged children in situations of social risk. This second goal has special implications because, to the best of our knowledge, this is the first study to examine the measurement invariance of the BRIEF-2 across genders, and in a sample with these specific characteristics. Through the validation, we intend to confirm the three-factor/nine-scale structure of the instrument and its measurement invariance across genders, examine its internal consistency, and explore possible gender differences on the BRIEF-2 scales and indexes.

Method

Participants

Participants in the study were a total of 168 students from 6 to 15 years old (M = 10.08; *SD* = 2.09), of whom 87 (51.8%) were girls. By age group, 47 students (28%) were from 6 to 8 years old; 76 (45.2%) were from 9 to 11 years old; 42 (25%) were from 12 to 14 years old; and 3 (1.8%) were 15 years old. An incidental sampling procedure was used from a population that participates in three schools in Santo Domingo (Dominican Republic), two of which also include a modality for children in residential foster care due to conditions of vulnerability and risk (75%). In general, the participants in the present study stem from an environment with quite precarious socioeconomic conditions (65% have income below the country's official general poverty level and extreme poverty level, Carneiro and Sirtaine, 2017). The parents or caregivers of the participants mainly had a basic or primary (48%) or secondary (21%) educational level.

In the case of the caregivers in the protective installations, they generally had university studies in progress (88%). The caregivers completed the tests in the case of boys and girls who did not have constant contact with their original or biological family.

Measures

The Behavior Rating Inventory of Executive Function, Second Revision, Parent Form (BRIEF-2; Gioia et al., 2015; Maldonado et al., 2017). The BRIEF-2 is a rating scale that makes it possible to explore the executive functioning in the daily behavior of children from 5 to 18 years old, observed and reported by parents or caregivers. It is composed of 63 items that make up nine scales, grouped in three indexes in the following way: the Shift and Emotional Control scales in the ERI; the Inhibit and Self-Monitor scales in the BRI; and the Initiate, Working Memory, Plan/Organize, Task Monitor, and Organization of Materials scales in the CRI. These three indexes make up the GEC. The parents have to indicate how often (never, sometimes, and often) the 63 behaviors have been a problem in the past six months. The participants completed the Spanish adaptation (Maldonado et al., 2017). The internal consistency values of the scales ranged between .66 (Initiative) and .87 (Emotional Control). For the indexes, all the values were above .85 (Maldonado et al., 2017).

Data analyses

Since we used the scale scores (or item parcels) to examine the dimensional structure and measurement invariance of the BRIEF-2, we first tested the dimensionality for each of the scales by means of confirmatory factor analysis (CFA). According to Little, Cunningham, Shahar, and Widaman (2002), if the scale to be parceled does not represent a unidimensional construct, parcels of items could mask multidimensionality. Once the dimensionality of each scales was tested, the descriptive statistics for the BRIEF-2 scales were calculated.

Second, we then conducted CFAs to assess the fit of the five potential BRIEF-2 models: two models with one dimension and eight or nine scales, respectively (Fournet et al., 2015), two models with two factors and eight or nine scales, respectively (Egeland and Fallmyr, 2010; Gioia, Isquith, Retzlaff et al., 2002), and one model with three factors and nine scales (Egeland and Fallmyr, 2010; Fournet et al., 2015; Gioia et al., 2015; Maldonado et al., 2017).

Third, in order to study the MI of the resulting baseline model across gender, successive multi-group CFAs were conducted. We tested three levels of group invariance: configural, metric, and strong invariance. For the configural invariance, manifest indicators were constrained to load on the same factor across groups. Next, factor loadings were constrained to be equal across groups (weak invariance). Finally, equality constraints on the loadings and intercepts were set to test the strong invariance. The confirmation of strong invariance allows the comparison of the latent factors (BRIEF-2 indexes) means in the two groups. Likewise, to examine the possible gender differences on the BRIEF-2 scales, analysis of variance (ANOVA) was used ($p < .001$). Effect sizes were expressed using η_p^2 . Finally, Cronbach's alphas for the BRIEF-2 scales and indexes were calculated.

All the analyses were performed using IBM SPSS Statistics 22.0 (SPSS, 2013) and Mplus 7.0 (Muthén and Muthén, 2012). Due to the continuous nature of the data, robust maximum likelihood (MLR) was used. As criteria for model fit, we report the chi-square test. the comparative fit index (CFI), the Tucker–Lewis index, the RMSEA, and the standardized root mean residual (SRMR). Hu and Bentler (1995) suggested that RMSEA should be ≤.08 for a reasonable model fit, and CFI and TLI should be above .95, although any value ≥.90 tends to be considered acceptable. For the SRMR values, a value below .08 has been suggested to indicate adequate model fit. In testing the invariance hypothesis, changes in the model fit of CFI (ΔCFI) were used. A decrease (≥.01) in CFI revealed an unacceptable decrease in model fit (Cheung and Rensvold, 2002).

Results

Unidimensionality and descriptive statistics of the scale scores

In order to use the scale scores (item parcels or group of items), we first tested the unidimensionality of each BRIEF-2 scale. As Table 1 shows, the goodness-of-fit indices revealed that all the scales with the exception of Task Monitor and Monitor (consisting of the Task Monitor and Self Monitor items) met the unidimensional condition. For these scales, the CFI, TLI, and RMSEA were close to meet the cutoff values (Hu and Bentler, 1995).

Table 2 shows the descriptive statistics for the BRIEF-2 scale scores (mean, standard deviation, skewness, and kurtosis). As the table shows, all values of skewness and kurtosis were inside the −1 to 1 range.

Sources of validity evidence based on the internal structure of the BRIEF-2

Table 3 shows the goodness-of-fit indices for the five models tested. The three-factor model presented better goodness-of-fit indices than the other two competing models. The correlations between the three latent factors ranged from .88 (Behavioral–Cognitive Regulation) to .93 (Behavioral–Emotional Regulation).

Table 1. Goodness-of-fit statistics resulting for the unidimensionality of the different scales (n = 168).

Scales	χ^2	*df*	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR
Inhibit	20.60	20	.998	.998	.013	.000 to .068	.031
Shift	30.99	20	.953	.934	.057	.000 to .095	.045
Emotional control	40.91***	20	.963	.945	.061	.012 to .099	.062
Initiate	8.73	5	.930	.861	.067	.000 to .138	.043
Working Memory	28.88	20	.964	.950	.051	.000 to .090	.044
Plan/Organize	21.28	20	.994	.991	.019	.000 to .070	.045
Organization of Materials	5.15	5	.998	.997	.013	.000 to .109	.029
Monitor	77.39***	35	.884	.847	.085	.059 to .111	.063
Self-Monitor	4.34	2	.969	.906	.083	.000 to .193	.028
Task Monitor	40.63***	9	.878	.860	.084	.052 to .137	.065

Note. CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual. ***p < .001.

Table 2. Descriptive statistics for the parent form of the BRIEF-2 (n = 168).

	Total		Girls		Boys					
Scales	M	*SD*	M	*SD*	M	*SD*	Skewness	Kurtosis	Range	Alpha
Inhibit	13.42	4.08	12.97	4.21	13.90	3.91	.45	−.73	8–23	.88
Shift	13.37	3.39	12.75	3.34	14.03	3.33	.35	−.55	8–23	.79
Emotional control	14.45	3.75	14.11	3.96	14.81	3.55	.19	−.49	8–24	.72
Initiate	8.56	2.34	8.12	2.43	9.03	2.16	.28	−.66	5–15	.67
Working Memory	14.06	3.66	13.85	3.64	14.28	3.70	.20	−.91	8–22	.81
Plan/Organize	13.98	3.42	13.60	3.33	14.36	3.48	.24	−.67	8–23	.76
Organization of Materials	10.34	2.95	9.87	2.97	10.84	2.87	.37	−.47	6–18	.79
Monitor	14.38	4.09	13.89	4.61	14.89	3.40	.11	−.57	9–26	.79
Self-Monitor	7.42	3.01	7.27	3.75	7.57	1.92	.07	−.48	4–12	.67
Task Monitor	8.71	2.21	8.34	2.13	9.10	2.23	.28	−.75	6–18	.66

Table 3. Goodness-of-fit statistics resulting for the different models tested (n = 168).

Model	χ^2	*df*	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR
1a One-factor, 8 scales	69.60**	20	.944	.922	.122	.091, .153	.036
1b One-factor, 9 scales	81.61**	27	.945	.927	.110	.083, .137	.035
2a Two factors, 8 scales	54.81**	19	.960	.941	.106	.074, .139	.036
2b Two factors, 9 scales	72.23**	26	.954	.936	.103	.075, .131	.036
3 Three factors, 9 scales	51.32*	24	.973	.959	.082	.051, .113	.029

Note. CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual. **p < .01. **p < .001.

Measurement invariance of the BRIEF-2 across gender

Given that the three-factor model showed the best fit, we tested the measurement equivalence of this model across gender. Prior to the analysis of measurement invariance, we tested whether this model showed a reasonably good fit to the data in each group separately (male and female). The results are shown in Table 4. The goodness-of-fit indices obtained for both gender groups in terms of CFI, TLI, and SRMR were satisfactory, whereas RMSEA exceeded the cutoff values (Hu and Bentler, 1995). The configural invariance model, where no equality constraints were imposed, showed an adequate fit to the data, again with the exception of the RMSEA value. Then we tested weak invariance for the two groups. The goodness-of-fit indices were all acceptable, with the exception of RMSEA. The ΔCFI between the configural and weak models was <.01, indicating that the hypothesis of weak invariance was tenable. Subsequently, strong MI was tested, where intercepts and factor loadings were constrained to be equal across

Table 4. Goodness-of-fit indexes of measurement invariance across gender for BRIEF-2 three-factor model.

Model	χ^2	*df*	CFI	TLI	RMSEA	RMSEA 90% CI	SRMR	ΔCFI
Gender								
Females (n = 87)	48.13*	24	.953	.929	.108	.062, .151	.043	
Males (n = 81)	45.88*	24	.960	.939	.106	.058, .152	.035	
Multiple group								
Configural	94.24**	48	.956	.934	.107	.075, .139	.039	
Weak	99.45**	57	.960	.949	.094	.062, .124	.059	.004
Strong	116.41**	66	.952	.948	.095	.066, .123	.089	.008

Note. CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual. *p < .01. **p < .001.

groups. Again, the RMSEA value was slightly above the cutoff values. The ΔCFI between the strong and weak invariance models was below .01, indicating that strong invariance was supported, based on the recommendations of Cheung and Rensvold (2002). Hence, overall the results support configural, weak, and strong invariance of the three-factor model of the BRIEF-2 across genders.

Gender differences of the BRIEF-2 index and scale scores

After confirming the measurement invariance of the BRIEF-2 across gender, we are able to examine the differences in the means on the indexes and scales. For the three index scores, latent mean differences across genders were estimated, fixing the latent mean values to zero in males (Marsh et al., 2014). For comparisons of the gender groups on the latent means, statistical significance was based on the *z*-statistic. Boys scored higher than girls in all three dimensions [BR (.472; $p = .069$), ER (.497; $p < .059$), CR (.462; $p < .063$)], although these differences were not statistically significant.

The ANOVA revealed that boys obtained higher scores than girls on all the nine scales, with statistically significant differences on: Shift ($M_{boys} = 12.75$; $SD_{boys} = 3.34$; $M_{girls} = 14.03$; $SD_{girls} = 3.33$; $F = 6.17$; $p < .05$; $\eta_p^2 = .036$); Initiate ($M_{boys} = 8.12$; $SD_{boys} = 2.43$; $M_{girls} = 9.03$; $SD_{girls} = 2.16$; $F = 6.60$; $p < .05$; $\eta_p^2 = .038$); Task-Monitor ($M_{boys} = 8.34$; $SD_{boys} = 2.13$; $M_{girls} = 14.37$; $SD_{girls} = 3.49$; $F = 5.08$; $p < .05$; $\eta_p^2 = .030$); and Organization of Materials ($M_{boys} = 9.87$; $SD_{boys} = 2.97$; $M_{girls} = 10.84$; $SD_{girls} = 2.87$; $F = 4.59$; $p < .05$; $\eta_p^2 = .027$).

Reliability estimation of the BRIEF-2 index and scale scores

The Cronbach's alpha coefficients for the index scores ranged from .71 (BR) to .91 (CR). The coefficient alpha values for the scales ranged from .66 (Initiate, five items) to .88 (Inhibit, eight items) (see Table 1).

Discussion

The BRIEF-2 is a widely used test in the reported assessment of the executive functioning in children from 5 to 18 years old. This study has two key objectives, first, to analyze the dimensional structure of the BRIEF-2 in its versions for parents or caregivers, and second, to investigate the measurement invariance across gender in a sample of primary school-aged children at risk of social exclusion.

With regard to the former, the BRIEF-2 structure with three factors and nine scales was confirmed (Gioia et al., 2015, Maldonado et al., 2017), and this structure has also been confirmed in other previous studies with the original BRIEF (e.g., Egeland and Fallmyr, 2010; Fournet et al., 2015; Gioia, Isquith, Kenworthy et al., 2002). This model is better than the models with one or two factors and eight or nine scales, respectively. The model with three factors and nine scales groups multidimensional aspects at behavioral (Inhibit and Self-Monitor), emotional (Shift and Emotional Control), and cognitive (Initiate, Working Memory, Plan/Organize, Task Monitor, and Organization of Materials scales) levels.

It is important to point out that this model also presents higher than optimal RMSEA values, a result that has been observed in the majority of studies that have examined the structure of the BRIEF, in both its first and second versions (Fournet et al., 2015; Gioia, Isquith, Retzlaff et al., 2002; Pérez-Salas, Ramos, Oliva, & Ortega, 2016; Maldonado et al., 2017; Peters et al., 2012). Recently, some authors (e.g., Sun et al., 2018) have mentioned the close relationships between some of the scales from different factors (e.g., between Inhibit and Working Memory) as a potential explanation for the larger RMSEA values.

On the other hand, it is also important to note that prior to the analysis of the dimensional structure of the BRIEF-2, we tested the unidimensionality of the different scales (Little et al., 2002) and found that two of the scales did not meet this criteria: the Monitor and the Task Monitor scales. This result might also partly explain the larger RMSEA values throughout all the examined models. To the best of our knowledge, no previous study on the dimensional structure of the BRIEF-2 first examined the advisability of using item parcels, and so we are unable to compare or contrast our results. Further studies should take this aspect into consideration.

In addition, the three-factor and nine-scale structure had acceptable psychometric properties in our sample, with acceptable to good scale score reliability and adequate construct validity, as highlighted by the correlations between the dimensions. The internal consistency values of the scales obtained in our study (range .66 to .88) are almost identical to those obtained in the Spanish adaptation (i.e., range .66 to .87; Maldonado et al., 2017), and lower than those found by Gioia et al. (2015) (range .80 to .91). In this regard, it should be pointed out that the internal consistency of the BR index obtained in our study (.71) was lower than the one obtained by Maldonado et al. (2017) (.88) and Gioia et al. (2015) (.90).

The second aim of our study was to examine MI by gender for the three-factor structure of the BRIEF-2, in order to further validate the measure and subsequently screen for gender differences in executive functioning. As far as we know, no previous studies have analyzed the MI of the BRIEF-2 across gender, and only Fournet et al. (2015) did so with the original BRIEF. In our study, the results supported the MI of the three-factor model across genders in our sample of primary school children at risk of social exclusion. Thus, the number of factors and their loading patterns (configural invariance), the factor loadings (weak invariance), and the item intercepts (strong invariance) were the same for boys and girls. We then examined gender differences in the nine scales and found that boys scored higher than girls on all nine scales, showing greater executive difficulties. However, these differences were only statistically different for the Shift, Task-Monitor, Initiate, and Organization of Materials scales.

When we examined the latent means, we also found that boys scored higher than girls on the three indexes, but these differences were not statistically significant. These results are partly consistent with those obtained by Fournet et al. (2015), the only study that has analyzed the MI of the BRIEF across gender. They found greater executive difficulties in boys, who obtained significantly higher scores on the Inhibit, Plan/Organize, Organization of Materials, and Monitor scales. Other studies carried out with the BRIEF (e.g., Donders et al., 2010) did not find gender differences. However, Maldonado (2016) obtained gender differences, with greater difficulties in boys, on several variables (Inhibit, Initiate, Working Memory, Plan, Monitor, and the

Metacognitive Index). Huizinga and Smidts (2011) also observed higher scores in boys on the Shift scale and on all the scales of the MI index (except the Organization of Materials scale). This previous research and our study findings suggest the need to continue to investigate gender differences on both the scales and indexes of the BRIEF-2.

This study has some limitations. First, it was carried out with a non representative sample with some specific characteristics (population in the Dominican Republic in situation of vulnerability), so that the results are not generalizable to the rest of the population. In addition, few studies are carried out in the Latin American context, and so the majority of the referents come from other societies with quite diverse social and contextual characteristics. On the other hand, although age is a relevant variable in the study of executive functioning, we could not examine the measurement invariance of the BRIEF-2 across age due to our sample size. Finally, the cross-sectional nature of this study kept us from examining the stability and evolution of the three-factor structure of the BRIEF-2 over time.

Future studies should continue to investigate the internal structure of the BRIEF-2 in other populations and cultures, etc., and evaluate its structure across age and over time (longitudinal measurement invariance). Studies of this type would make it possible to investigate developmental changes in EF in school-aged children.

The results obtained in the present study provide new evidence about the internal validity of the BRIEF-2 for parents and caregivers in Spanish in a primary school student population in situations of vulnerability. The test presents adequate psychometric properties, which means that the parent and caregiver version of this instrument is useful for assessing the executive functioning based on reports of behaviors observed in the children, and it makes it possible to detect possible situations of risk in their development in order to create appropriate interventions. It is important to bear in mind that, although the assessed population speaks Spanish, they live in the Dominican Republic (Latin American context), which is culturally different from Spain (European context). Validating the structure of the BRIEF-2 in an under-studied population with very special characteristics makes this instrument more susceptible to being applied in culturally diverse populations. Thus, the results of the present study contribute to understanding the executive functions using a multidimensional model of related behavioral, cognitive, and emotional components, as proposed in the scientific community, although this should continue to be explored in other populations and contexts.

ORCID

Beatriz Lucas-Molina http://orcid.org/0000-0003-4370-4799

References

Abundis, A. (2014). *Attention and executive function: Development and influence of socioenvironmental factors*. (Dissertation thesis). University of Granada: Spain.

Anderson, V. (2002). Executive function in children: Introduction. *Child Neuropsychology*, *8*(2), 69–70.

Barkley, R. A. (2012). *Executive functions: What they are, how they work, and why they evolved*. New York, NY: Guilford Press.

Berthelsen, D., Hayes, N., White, S. L., and Williams, K. E. (2017). Executive function in adolescence: Associations with child and family risk factors and self-regulation in early childhood. *Frontiers in Psychology*, *8*(903), 1–14.

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., and Domitrovich, C. E. (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the head start REDI program. *Development and Psychopathology*, *20*(3), 821–843.

Blair, C., and Raver, C. C. (2012). Individual development and evolution: Experiential canalization of self-regulation. *Developmental Psychology*, *48*(3), 647–657.

Carneiro, F. G., and Sirtaine, S. (Eds.). (2017). *When growth is not enough: explaining the rigidity of poverty in the dominican republic*. Washington, DC: World Bank Publications. doi:10.1596/978-1-4648-1036-7

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

Dodzik, P. (2017). Behavior Rating Inventory of Executive Function, Second Edition Gerard A. Gioia, Peter K. Isquith, Steven C. Guy, and Lauren Kenworthy. *Journal of Pediatric Neuropsychology*, *3*(3–4), 227–231.

Donders, J., DenBraber, D., and Vos, L. (2010). Construct and criterion validity of the behaviour rating inventory of executive function (BRIEF) in children referred for neuropsychological assessment after pediatric traumatic brain injury. *Journal of Neuropsychology*, *4*(2), 197–209.

Egeland, J., and Fallmyr, Ø. (2010). Confirmatory factor analysis of the behavior rating inventory of executive function (BRIEF): Support for a distinction between emotional and behavioral regulation. *Child Neuropsychology*, *16*(4), 326–337.

Fournet, N., Roulin, J. L., Monnier, C., Atzeni, T., Cosnefroy, O., Le Gall, D., and Roy, A. (2015). Multigroup confirmatory factor analysis and structural invariance with age of the behavior rating inventory of executive function (BRIEF) - French version. *Child Neuropsychology*, *21*(3), 379–398.

Gioia, G. A., Isquith, P. K., Guy, S. C., and Kenworthy, L. (2000). *Behavior rating inventory of executive function professional manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Gioia, G. A., Isquith, P. K., Guy, S. C., and Kenworthy, L. (2015). *BRIEF-2. Behavior rating inventory of executive function* (2nd ed.). Lutz, FL: Psychological Assessment Resources, Inc.

Gioia, G. A., Isquith, P. K., Kenworthy, L., and Barton, R. M. (2002). Profiles of everyday executive function in acquired and developmental disorders. *Child Neuropsychology*, *8*(2), 121–137.

Gioia, G. A., Isquith, P. K., Retzlaff, P. D., and Espy, K. A. (2002). Confirmatory factor analysis of the behavior rating inventory of executive function (BRIEF) in a clinical sample. *Child Neuropsychology*, *8*(4), 249–257.

Hu, L., and Bentler, P. M. (1995). Evaluating model fit. *In* R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 76–99). London: Sage.

Huizinga, M., and Smidts, D. P. (2011). Age-related changes in executive function: A normative study with the Dutch version of the behavior rating inventory of executive function (BRIEF). *Child Neuropsychology*, *17*(1), 51–66.

Jacobson, L. A., Pritchard, A. E., Koriakin, T. A., Jones, K. E., and Mahone, E. M. (2016). Initial examination of the BRIEF-2 in clinically referred children with and without ADHD symptoms. *Journal of Attention Disorders*, Online first. doi: 10.1177/1087054716663632

Little, T. D., Cunningham, W. A., Shahar, G., and Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*, 151–173.

Lyons Usher, A. M., Leon, S. C., Stanford, L. D., Holmbeck, G. N., and Bryant, F. B. (2016). Confirmatory factor analysis of the behavior rating inventory of executive functioning (BRIEF) in children and adolescents with ADHD. *Child Neuropsychology*, *22*(8), 907–918.

Maldonado, M. J. (2016). *Adaptación del BRIEF (Behavior Rating Inventory of Executive Function) a población española y su utilidad para el diagnóstico del trastorno por déficit de*

atención-hiperactividad subtipos inatento y combinado (Doctoral dissertation, Complutense University of Madrid, Spain). [Adaptation of the BRIEF (Behavior Rating Inventory of Executive Function) to the Spanish population and its usefulness for the diagnosis of attention deficit hyperactivity disorder, inattentive and combined subtypes]. Retrieved from http://eprints.ucm.es/37563/1/T37182.pdf

Maldonado, M. J., Fournier, C., Martínez, R., González, J., Espejo-Saavedra, J. M., and Santamaría, P. (2017). *BRIEF-2. Evaluación Conductual de la Función Ejecutiva [BRIEF-2. Behavioral assessment of the executive functioning]*. Madrid: TEA Ediciones.

Marsh, H. W, Morin, A. J, Parker, P. D, and Kaur, G. (2014). Exploratory structural equation modeling: an integration of the best features of exploratory and confirmatory factor analysis. *Annual Review Of Clinical Psychology*, *10*, 85-110. doi: 10.1146/annurev-clinpsy-032813-153700

Masten, A. S., Herbers, J. E., Desjardins, C. D., Cutuli, J. J., McCormick, C. M., Sapienza, J. K., … Zelazo, P. D. (2012). Executive function skills and school success in young children experiencing homelessness. *Educational Researcher*, *41*(9), 375–384.

McDermott, J. M, Westerlund, A, Zeanah, C. H, Nelson, C. A, and Fox, N. A. (2012). Early adversity and neural correlates of executive function: implications for academic adjustment. *Developmental Cognitive Neuroscience*, *2*, S59-S66. doi:10.1016/j.dcn.2011.09.008

Memisevic, H. (2015). Factor structure of behavior rating inventory of executive functions in children with intellectual disability. *Acta Neuropsychologica*, *13*(2), 137–144.

Muthén, L. K., and Muthén, B. O. (2012). *MPlus: Statistical analysis with latent variables. User's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Obradović, J. (2016). Physiological responsivity and executive functioning: Implications for adaptation and resilience in early childhood. *Child Development Perspectives*, *10*(1), 65–70.

Pérez-Salas, C. P., Ramos, C., Oliva, K., and Ortega, A. (2016). Bifactor modeling of the behavior rating inventory of executive function (BRIEF) in a Chilean sample. *Perceptual and Motor Skills*, *122*(3), 757–776.

Peters, C., Algina, J., Smith, S. W., and Daunic, A. P. (2012). Factorial validity of the behavior rating inventory of executive function (BRIEF)-teacher form. *Child Neuropsychology*, *18*(2), 168–181.

Raver, C. C., Blair, C., and Willoughby, M. (2013). Poverty as a predictor of 4-year-olds' executive function: New perspectives on models of differential susceptibility. *Developmental Psychology*, *49*(2), 292–304.

Roos, L. E., Kim, H. K., Schnabler, S., and Fisher, P. A. (2016). Children's executive function in a CPS-involved sample: Effects of cumulative adversity and specific types of adversity. *Children and Youth Services Review*, *71*, 184–190.

Slick, D. J, Lautzenhiser, A, Sherman, E M, and Eyrl, K. (2006). Frequency of scale elevations and factor structure of the behavior rating inventory of executive function (brief) in children and adolescents with intractable epilepsy. *Child Neuropsychology*, *12*(3), 181–189.

SPSS. (2013). *IBM SPSS statistics for windows. Version 22.0*. Armonk, NY: IBM Corp.

Stelzer, F., Mazzoni, C. C., and Cervigni, M. A. (2013). Cognitive models of executive functions development. Methodological limitations and theoretical challenges. *Anales De Psicología/Annals of Psychology*, *30*(1), 329–336.

Sun, X., Wu, Z., Cao, Q., Qian, Y., Liu, Y., Yang, B., … Wang, Y. (2018). Genetic variant for behavioral regulation factor of executive function and its possible brain mechanism in attention deficit hyperactivity disorder. *Scientific Reports*, *8*(1), 7620.

Zelazo, P., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., … Carlson, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, *68*(3), I–151.

Zelazo, P. D., and Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*(4), 354–360.