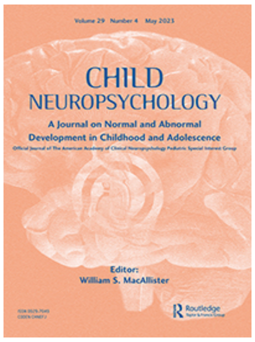


Child Neuropsychology

A Journal on Normal and Abnormal Development in Childhood and Adolescence

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ncny20

Factor structure and measurement invariance of the BRIEF2 Parent Form across gender in a Portuguese sample

Octávio Moura, Cristina P. Albuquerque, Marcelino Pereira, Sofia Major, Ana Filipa Lopes, Manuela Vilar, Maria João Seabra-Santos & Mário R. Simões

To cite this article: Octávio Moura, Cristina P. Albuquerque, Marcelino Pereira, Sofia Major, Ana Filipa Lopes, Manuela Vilar, Maria João Seabra-Santos & Mário R. Simões (2023) Factor structure and measurement invariance of the BRIEF2 Parent Form across gender in a Portuguese sample, Child Neuropsychology, 29:4, 628-643, DOI: 10.1080/09297049.2022.2105313

To link to this article: https://doi.org/10.1080/09297049.2022.2105313

Published online: 01 Aug 2022.

CHILD NEUROPSYCHOLOGY
2023, VOL. 29, NO. 4, 628–643
https://doi.org/10.1080/09297049.2022.2105313

Factor structure and measurement invariance of the BRIEF2 Parent Form across gender in a Portuguese sample

Octávio Moura[a], Cristina P. Albuquerque[a], Marcelino Pereira[a], Sofia Major[a,b], Ana Filipa Lopes[a], Manuela Vilar[a], Maria João Seabra-Santos[a] and Mário R. Simões[a]

[a]University of Coimbra, Center for Research in Neuropsychology and Cognitive and Behavioural Intervention, Faculty of Psychology and Educational Sciences, Coimbra, Portugal; [b]University of the Azores, Azores, Portugal

ABSTRACT

This study aimed to investigate the factor structure and the measurement invariance across gender of the BRIEF2 Parent Form in Portuguese typically developing children. Participants were 700 typically developing children ($n = 352$ girls and $n = 348$ boys) aged 6–16 years. Confirmatory factor analysis was used to test five competing factor models. Consistent with the BRIEF2 original dimensional structure, the three-factor model demonstrated the most adequate fit to the data. The measurement invariance of the three-factor model across gender was supported (configural, metric, and partial scalar invariance). Overall, the BRIEF2 Parent Form showed adequate psychometric properties, suggesting that it is a useful instrument to assess everyday executive functioning based on reports of behaviors observed by parents in healthy Portuguese children.

ARTICLE HISTORY
Received 15 May 2022
Accepted 19 July 2022

KEYWORDS
BRIEF2; executive functions; confirmatory factor analysis; measurement invariance

Introduction

Executive functions (EF) are a complex, multidimensional construct that encompasses cognitive processes required for conscious, top-down control of behaviors, thoughts, and emotions. It refers to metacognitive capacities that allow an individual to perceive stimuli, respond adaptively, establish goals, flexibly change actions, monitor results and respond in an integrated way (Baron, 2018).

Despite its wide acceptance, there is no consensus about the executive components involved (e.g., single construct or multiple interrelated functions; the number of processes and factors/domains). That is, EF is an umbrella term comprising a wide range of higher-order cognitive processes, such as flexibility/shifting, inhibition, planning, problem-solving, working memory, updating, and initiation, among others (Chan et al., 2008; Friedman & Miyake, 2017; Jurado & Rosselli, 2007; Wasserman & Wasserman, 2013). There is also considerable behavioral and neural evidence that EF varies along a continuum from "hot" (used in emotional and motivational contexts) to

CONTACT Octávio Moura octaviomoura@gmail.com Faculty of Psychology and Educational Sciences, Center for Research in Neuropsychology and Cognitive and Behavioural Intervention, University of Coimbra Rua do Colégio Novo, Coimbra 3000-115, Portugal

"cool" EF (used in relatively neutral emotionally and motivational contexts), as well as a dynamic interaction between more bottom-up (e.g., limbic) and more top-down (cortical) influences on information processing and behavior (Zelazo & Carlson, 2012; Zelazo, 2020).

Their importance is evident in the fact that they are central to school readiness, early school achievement, and academic performance throughout childhood and adolescence (Blair & Raver, 2015; Muñoz & Filippetti, 2021). They may also protect against the academic risks associated with poverty (Masten et al., 2012), contribute to social competence and social skills (Riggs et al., 2004), and some EF are relevant cognitive phenotypes for a variety of neurodevelopmental disorders (Bednarz et al., 2020; Moura et al., 2014; Robinson et al., 2009; Willcutt et al., 2005).

Therefore, it is essential to assess EF, with the most frequently applied rating scale for that purpose being the Behavior Rating Inventory of Executive Function (BRIEF), either in its original edition (Gioia et al., 2000) or in its second and revised edition (BRIEF2; Gioia et al., 2015). Indeed, the BRIEF and the BRIEF2 are widely used in clinical, educational, and research settings, and their popularity may be attributed to a variety of reasons, among others, with their prominent features being: (i) the ability to discriminate a large range of disorders from normal functioning across age groups (Gioia et al., 2000, 2015; Granader et al., 2014); (ii) the ability to identify distinct EF deficits in specific disorders (e.g., attention-deficit/hyperactivity disorder (ADHD) and oppositional defiant disorder; (Halvorsen et al., 2019); (iii) the ability to assess several EF in an ecologically valid way, i.e., as they manifest themselves in daily life; (iv) sensitivity with respect to treatment interventions (Gioia et al., 2015); (v) responses coming from different information sources, thus allowing for a more global assessment of EF; (vi) ease and speed of administration.

BRIEF and BRIEF2 factorial structure

The BRIEF was published in 2000 (Gioia et al., 2000) and it was the first published rating scale to measure everyday behaviors associated with EF in children and adolescents aged 5–18 years. It was designed to be completed by parents and teachers, although a few years later an adolescents' self-report version was also developed (Guy et al., 2004).

The BRIEF comprised 86 items distributed across eight scales: Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor. An exploratory factor analysis of these eight scale scores revealed a two-factor solution both in the BRIEF standardization sample and in a mixed clinical sample (Gioia et al., 2000). One of the factors was designated Behavioral Regulation Index (BRI) and incorporated the Inhibit, Shift, and Emotional Control scales, while the other factor was designated the Metacognition Index (MI) and comprised the remaining five scales. Two years later, Gioia et al. (2002) performed a confirmatory factor analysis (CFA) of the Parent Form of the BRIEF in a mixed clinical sample of children. They tested a revised nine-scale BRIEF configuration that separated the Monitor scale into a Self-Monitor and a Task Monitor scale. A three-factor solution best fitted the data when compared with one-, two- or four-factor models: the BRI now consisted of the Inhibit and Self-Monitor scales, and a new Emotion Regulation Index (ERI) included the Shift and Emotional Control scales, whereas the MI did not undergo any alterations.

Since the Gioia et al. (2002) study, several analyses of the factor structure of the BRIEF Parent Form have been conducted with different samples and using either the original eight-scale composition, the nine-scale composition, or both. A two-factor model (BRI and MI) of the eight scales showed an adequate fit to the data in clinical samples of children with epilepsy (Slick et al., 2006), traumatic brain injury (Donders et al., 2010) and ADHD (Usher et al., 2016), as well as in the Dutch standardization sample of the BRIEF (Huizinga & Smidts, 2011). Using data gathered from parents of children with heterogeneous developmental disorders, Halvorsen et al. (2019) found that a three-factor model (BRI, ERI, and MI) of the nine BRIEF scales had adequate fit indices, although the root mean square error of approximation (RMSEA) value was marginally acceptable.

Studies of CFAs of both the eight and nine scales of the BRIEF Parent Form were also carried out in a clinical sample (Granader et al., 2014), in a mixed sample of healthy and clinically referred children (Egeland & Fallmyr, 2010), and in samples of typically developing children (TDC; Fournet et al., 2015; Granader et al., 2014). Granader et al. (2014) assessed three-, eight- and nine-factor models of the BRIEF and found that the nine-factor model met thresholds for goodness-of-fit in the sample of TDC, while in the clinical sample of children with autism spectrum disorder, none of the models met goodness-of-fit criteria. In a mixed sample, Egeland and Fallmyr (2010) observed the superiority of the three-factor model based on the nine scales, although some of the fit indices (adjusted goodness of fit index and RMSEA) were below satisfactory levels. In the French standardization sample, a three-factor model and a two-factor model, based on a nine-scale structure, showed a good fit (Fournet et al., 2015).

In 2015, the BRIEF was revised, and this second edition has become commonly referred to as the BRIEF2. The number of items was reduced to 63, and the nine-scale configuration was implemented. A CFA in the standardization sample and in a combined clinical sample supported the three-factor model of BRI, ERI, and Cognitive Regulation Index (CRI; renamed from the MI) (Gioia et al., 2015). From 2015 to the present, few independent factor analysis studies of the BRIEF2 Parent Form have been carried out; most studies (five) used clinical or at-risk samples and just two studies were performed in samples of TDC. CFAs were conducted on children living in conditions of risk and social vulnerability in the Dominican Republic (Jiménez & Lucas-Molina, 2019), in samples of Chinese (Shum et al., 2021) or Iranian (Parhoon et al., 2021) children with ADHD, and in samples of heterogeneous clinically referred American children (Jacobson et al., 2020; Lace et al., 2021). Most of these studies found support for a three-factor model with nine scales, although the RMSEA value was slightly above the cutoff in the Jiménez and Lucas-Molina study (Jiménez & Lucas-Molina, 2019). In contrast, Lace et al. (2021) found that three-factor models provided a poor fit and that a two-factor structure provided the best fit. It is also worth mentioning that an exploratory factor analysis suggested that the BRIEF2 might have five to six scales rather than nine (Jacobson et al., 2020). The CFAs of the BRIEF2 Parent Form were also carried out in samples of Spanish (Gioia et al., 2017) and Chilean TDC (Muñoz & Filippetti, 2021). Both studies indicated that the three-factor model best fit the data, although RMSEA values were higher than the recommended ones.

In conclusion, the evidence concerning the factor structure of the BRIEF and BRIEF2 is not entirely consistent. In addition, most of the studies with the BRIEF2 were carried out in clinical samples, and only a few concerned TDC (Gioia et al., 2015, 2017; Muñoz &

Filippetti, 2021). Furthermore, to our knowledge, only one study has examined the measurement invariance of the BRIEF2 across gender (Jiménez & Lucas-Molina, 2019) in a sample of socially vulnerable children. It is surprising because measurement invariance should be tested before performing group differences (e.g., gender, age, clinical samples) in order to analyze if the factor structure operates equivalently across different populations (Sideridis et al., 2015; Wicherts, 2016). Indeed, some studies have found significant differences between boys and girls (in favor of girls; e.g., Fournet et al., 2015; Jiménez & Lucas-Molina, 2019), and standardized data have been carried out separately for boys and girls (Gioia et al., 2000, 2015).

The present study

Even though Portuguese is the sixth most spoken language in the world with more than 200 million native speakers (Lewis et al., 2015), this is the first study that explores the factor structure and the measurement invariance of BRIEF2 Parent Form in European Portuguese-speaking children. Thus, this study aimed to investigate: (i) the factor structure of the BRIEF2 Parent Form (Portuguese version) in a sample of TDC since clinical samples have been predominant; and (ii) whether the BRIEF2 factor structure would operate equivalently across gender. Based on the existing literature, we hypothesized that the original three-correlated-factor model would be the most interpretable and parsimonious factor solution in a sample of Portuguese TDC aged 6 to 16 years old (Fournet et al., 2015; Gioia et al., 2015, 2017; Muñoz & Filippetti, 2021), and the factor structure would be equivalent across gender (Jiménez & Lucas-Molina, 2019).

Method

Participants

The participants were 700 Portuguese children and adolescents ($n = 352$ girls and $n = 348$ boys) from 6 to 16 years old ($M = 10.50$, $SD = 3.06$; 6y $n = 60$, 7y $n = 87$, 8y $n = 80$, 9y $n = 64$, 10y $n = 88$, 11y $n = 69$, 12y $n = 59$, 13y $n = 41$, 14y $n = 42$, 15y $n = 62$, and 16y $n = 48$), attending school from the first to eleventh grade. Boys and girls were matched for age (boys: $M = 10.51 \pm 3.01$; girls: $M = 10.49 \pm 3.12$; $t(698) = 0.111$, $p = .912$) and school grade (boys: $M = 5.20 \pm 2.96$; girls: $M = 5.24 \pm 3.07$; $t(698) = > -0.177$, $p = .860$). Participants were recruited from six public schools in urban (82%) and rural (18%) areas, close to the Portuguese distribution (Pordata, 2018). Most of the participants were from families of middle socioeconomic status (51%), with the rest being from lower (27.8%) or higher levels (21.2%). As for the educational level of the mothers of the children and adolescents being studied, the majority had a university degree (25.8% elementary, 30.4% secondary, and 43.8% university), whereas most of the fathers had only an elementary level of education (38.3% elementary, 30.4% secondary, and 31.3% university). Parents were invited to complete the BRIEF2, and most of the questionnaires were completed by the mother (81.6% from the mother, 15% from the father, and 3.4% from both parents).

Based on the parents' and teachers' data, children were excluded from the study if they had a diagnosis of neurodevelopmental disorders (e.g., ADHD, specific learning disorder, autism spectrum disorder), neurological disorders (e.g., epilepsy, traumatic brain injury), psychopathology (e.g., depressive disorders, anxiety disorders), disruptive, impulse-control, and conduct disorders (e.g., oppositional defiant disorder, conduct disorder), and special educational needs.

Measure

As previously mentioned, the BRIEF2 Parent Form (Gioia et al., 2015) consists of 63 items that are organized into three composite indexes and nine scales: (1) BRI: Inhibit, and Self-Monitor scales; (2) ERI: Shift, and Emotional Control scales; and (3) CRI: Initiate, Working Memory, Plan/Organize, Task-Monitor, and Organization of Materials scales. The BRIEF2 also has a unitary Global Executive Composite (GEC) that comprises the nine scales. In addition, it has three embedded validity scales (Inconsistency, Negativity, and Infrequency) assessing inconsistent, negative, and atypical response biases. Parents indicated how often their child has displayed a specific behavior in the last 6 months on a 3-point Likert scale (never, sometimes, and often), with higher scores indicating poorer EF.

Procedures

After approval by the Directorate-General for Education of the Portuguese Ministry of Education, the Ethics Commission of the Faculty of Psychology and Educational Sciences of the University of Coimbra (CEDI/FPCEUC: 5July 3 2021-07-2021), and each of the participants' school boards, parents were invited to enroll in the study. In each of the six schools, classrooms were randomly selected. All the parents of the selected classrooms were contacted by letter and invited to participate in the study (60% to 70% of the parents agreed to participate). Voluntary participation was requested from all parents, and the objectives of the study were fully explained. Written informed consent was obtained from the parents before the inclusion of participants in the study. Parents and teachers further provided information about the child's medical and educational history. No incentives (fees or extra credit) were offered in exchange for participation.

In accordance with the BRIEF2 Manual (Gioia et al., 2015), 35 participants were dropped out due to missing values (i.e., the total number of unanswered items was greater than 12, if two or more items of the same scale have a missing response, or if the scores of the validity scales [Inconsistency, Negativity, and Infrequency] were considered highly elevated [percentile >99]). Twenty-four children were eliminated based on the exclusion criteria (the presence of a neurodevelopmental or neurological disorder, psychopathology, disruptive behavior, and special educational needs).

Statistical analyses

Raw scores were used in all the statistical analyses. Descriptive statistics, correlation, reliability, and inferential analyses were conducted using IBM SPSS Statistics version 25.

To test the factor structure of the BRIEF2 Parents Form, a CFA was performed using IBM SPSS Amos 25 through the covariance matrix using maximum likelihood estimation. Hu and Bentler (1999) recommend a comparative fit index (CFI) of >.95, a standardized root mean square residual (SRMR) of <.08, and a RMSEA of <.06 to determine a good model fit. For the RMSEA, other cutoff values are also suggested: <.05 good fit, .05–.08 acceptable fit, .08–.10 mediocre fit, and >.10 poor fit (Byrne, 2010; MacCallum et al., 1996). The Akaike Information Criterion (AIC) was used to compare models, with smaller values representing a better fit. These traditional cutoff values should not be used as rules of thumb, and more stringent cutoff values are recommended for simple models and less stringent cutoff values are recommended for more complex models (Cheung & Rensvold, 2002; Marsh et al., 2004).

Measurement invariance was tested through the analysis of mean and covariance structures that encompassed a series of hierarchically ordered steps that began with the establishment of a baseline model for each group separately, followed by tests for increasingly more stringent levels of constrained equivalence across both groups: (i) *configural invariance*, no equality constraints were imposed on the parameters across the two groups; (ii) *metric invariance* ("weak factorial invariance"), factor loadings were constrained to be equal across groups; and (iii) *scalar invariance* ("strong factorial invariance"), factor loadings and intercepts were constrained to be equal across groups. It is commonly accepted that evidence for invariance is obtained if the multi-group model exhibits an adequate fit to the data, the χ^2 difference value ($\Delta\chi^2$) is not statistically significant ($p > .05$), and the CFI difference value (ΔCFI) is < −.010 (Byrne, 2010; Chen, 2007; Cheung & Rensvold, 2002).

Results

Descriptive statistics, correlation coefficients, and internal consistency

Skewness and kurtosis values were examined to determine the normality of the data distribution. As shown in Table 1, all of the scales showed skewness and kurtosis values <1, which suggested adequate distribution for maximum likelihood estimation (Byrne, 2010; Curran et al., 1996).

Table 1. Descriptive statistics and unidimensionality (CFA goodness-of-fit statistics).

	Descriptive statistics			Unidimensionality (CFA)		
	M (SD)	Skewness	Kurtosis	CFI	SRMR	RMSEA (90% CI)
Inhibit	11.89 (3.83)	0.836	0.547	.974	.027	.056 (.040–.073)
Self-Monitor	6.13 (1.83)	0.698	0.073	.992	.021	.054 (.028–.080)
Shift	11.69 (2.68)	0.798	0.556	.968	.035	.045 (.028–.062)
Emotional Control	11.62 (3.16)	0.878	0.505	.978	.027	.053 (.037–.069)
Initiate	7.54 (2.06)	0.599	−0.193	.993	.019	.042 (.000–.080)
Working Memory	12.11 (3.37)	0.805	0.125	.980	.030	.056 (.039–.073)
Plan/Organize	12.88 (3.23)	0.483	−0.162	.985	.026	.042 (.024–.060)
Task Monitor	8.40 (2.20)	0.469	−0.208	.995	.015	.045 (.000–.088)
Org. Materials	8.58 (2.24)	0.971	0.968	.969	.048	.081 (.058–.107)
BRI	18.01 (4.48)	0.828	0.662			
ERI	23.31 (5.17)	0.777	0.409			
CRI	49.53 (11.11)	0.626	−0.027			
GEC	90.73 (18.06)	0.597	0.033			

BRI = Behavior Regulation Index. ERI = Emotion Regulation Index. CRI = Cognitive Regulation Index. GEC = Global Executive Composite. CFI = Comparative Fit Index. SRMR = Standardized Root Mean Square Residual. RMSEA (90% CI) = Root Mean Square Error of Approximation (90% confidence interval).

Table 2. Pearson correlation coefficients and internal consistency.

	Correlation coefficients												
	2	3	4	5	6	7	8	9	BRI	ERI	CRI	GEC	α
1. Inhibit	.641	.431	.645	.418	.521	.491	.523	.483	.950	.611	.575	.756	.833
2. Self-Monitor		.456	.541	.479	.506	.512	.463	.424	.849	.526	.568	.719	.771
3. Shift			.555	.573	.499	.558	.376	.346	.484	.863	.557	.704	.728
4. Emotional Control				.388	.412	.476	.370	.383	.667	.899	.483	.712	.837
5. Initiate					.733	.740	.588	.508	.485	.536	.839	.781	.749
6. Working Memory						.763	.665	.626	.566	.506	.916	.840	.851
7. Plan/Organize							.638	.565	.551	.579	.896	.847	.807
8. Task Monitor								.520	.550	.417	.794	.739	.768
9. Org. Materials									.505	.409	.749	.696	.740
BRI										.653	.629	.814	.874
ERI											.584	.803	.855
CRI												.932	.937
GEC													.953

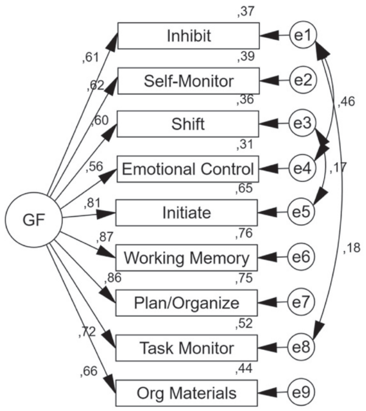
All correlation significant at $p < .001$. BRI = Behavior Regulation Index. ERI = Emotion Regulation Index. CRI = Cognitive Regulation Index. GEC = Global Executive Composite. The gray background color highlights the correlation coefficient between the scales of the same domain/index and between the scale and their domain/index.

Pearson correlations between the nine scales and the three composite indexes were computed. According to Cohen's guidelines (Cohen, 1988; small r = .1 , medium r = .3 , and large r = .5), medium to large correlation coefficients were observed between the scales. The nine scales showed large correlation coefficients with their composite index and GEC. Strong correlations between the three composite indexes ($r \geq .584$) and the GEC ($r \geq .855$) were also observed. The scales, composite indexes, and GEC showed acceptable internal consistency with Cronbach's α ranging from .728 to .953 (see Table 2).

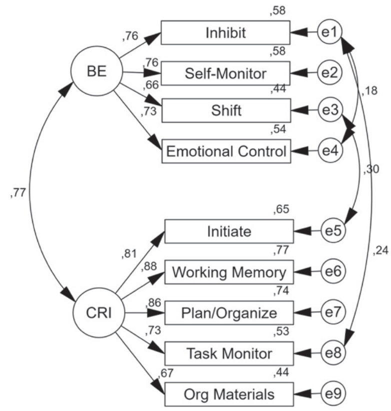
Confirmatory factor analysis

To perform a CFA at the scale level (i.e., scales were treated as continuous indicators), we first analyzed the unidimensionality of the nine BRIEF2 Parent Form scales. As shown in Table 1, the goodness-of-fit indices were adequate, suggesting that the scales were unidimensional.

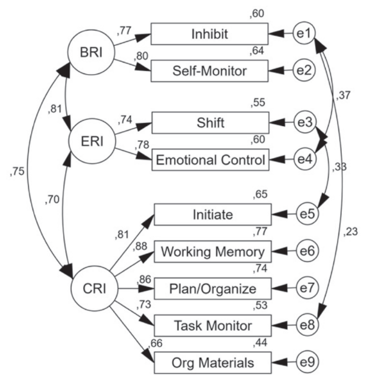
We tested five factor structures that are commonly used in factor analysis studies with BRIEF and BRIEF2 in different languages (e.g., Egeland & Fallmyr, 2010; Fournet et al., 2015; Gioia et al., 2002, 2015; Jiménez & Lucas-Molina, 2019; Lace et al., 2021; Muñoz & Filippetti, 2021; Parhoon et al., 2021; Pérez-Salas et al., 2016): (i) a one-factor model (a general executive factor in line with the view of EF as a unitary construct) with the nine scales; (ii) a two-correlated-factor model with the Inhibit, Self-Monitor, Shift, and Emotional Control scales in one factor (behavior/emotion factor) and the other five scales from the CRI in a second factor; (iii) the original three-correlated-factor model; (iv) a hierarchical three-factor model with the GEC (second-order factor), the three indexes (first-order factor), and the nine scales; and (v) a bifactor model where each scale loads onto its respective first-order factor (index) and simultaneously onto a general factor (see Figure 1).



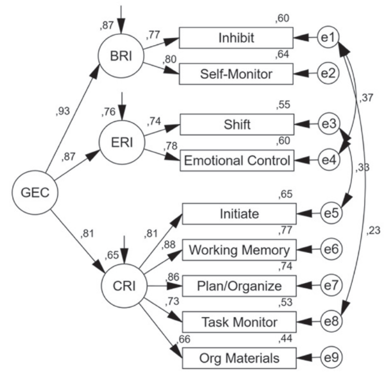
Model 1 – one-factor model
χ^2 = 230.147, CFI = .940, SRMR = .051, RMSEA (90% IC) = .114 (.101–.127), AIC = 272.147



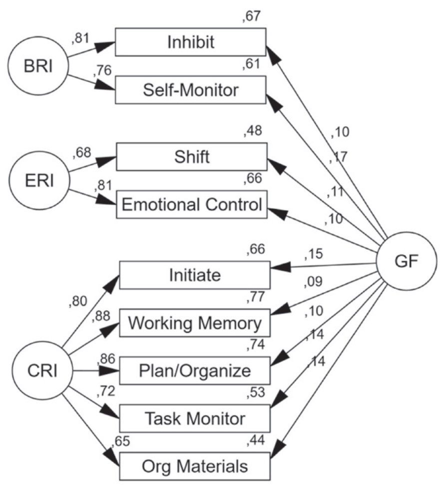
Model 2 – two-correlated-factor model
χ^2 = 158.866, CFI = .961, SRMR = .037, RMSEA (90% IC) = .094 (.081–.109), AIC = 202.866



Model 3 – three-correlated-factor model
χ^2 = 104.887, CFI = .976, SRMR = .030, RMSEA (90% IC) = .078 (.063–.093), AIC = 152.887



Model 4 – hierarchical three-factor model
χ^2 = 104.887, CFI = .976, SRMR = .030, RMSEA (90% IC) = .078 (.063–.093), AIC = 152.887



Model 5 – bifactor model
χ^2 = 226.564, CFI = .941, SRMR = .047, RMSEA (90% IC) = .116 (.102–.129), AIC = 270.564

Figure 1. BRIEF2 Parent Form – CFA standardized solution. BRI = Behavior Regulation Index. ERI = Emotion Regulation Index. CRI = Cognitive Regulation Index. GEC = Global Executive Composite. BE = Behavior/Emotion Factor. GF = General Factor. In model 5 (bifactor model), the error variances are not shown to reduce clutter.

BRI = Behavior Regulation Index. ERI = Emotion Regulation Index. CRI = Cognitive Regulation Index. GEC = Global Executive Composite. BE = Behavior/Emotion Factor. GF = General Factor. In model 5 (bifactor model), the error variances are not shown to reduce clutter.

A preliminary analysis of the modification indices for each of the factor models (except for the bifactor model) suggested the addition of three error covariances (between Inhibit – Emotional Control, Inhibit – Task Monitor, and Shift – Initiate). All the models have acceptable goodness-of-fit indices without error covariances (except the one-factor model), but the addition of these three error covariances led to an increase in models' fit. For example, the three-factor models (correlated and hierarchical) without the error covariances yielded a CFI =.941, SRMR =.047, RMSEA (90% IC) = .113 (.100–.127), and AIC = 268.735. The three-factor models (correlated and hierarchical) showed a better model fit than the other three competing models (see Figure 1). The correlated and hierarchical three-factor models revealed equal goodness-of-fit indices. The original three-correlated-factor model revealed adequate factor loadings and the three factors were highly correlated. The analysis of the average variance extracted showed evidence of convergent validity (AVE > .50) for all of the three factors (AVE = .62, .58, and .63, respectively).

Measurement invariance across gender

Given that the original three-correlated-factor model showed the best fit, we additionally performed a multiple-group analysis to evaluate whether the factor structure of the BRIEF2 Parent Form would be equivalent across gender.

Two CFAs were initially conducted for boys (CFI = .967, SRMR =.035, RMSEA =.091) and girls (CFI = .982, SRMR =.030, RMSEA =.065), yielding adequate models fit. After establishing the baseline model for each group, we tested for configural invariance. The configural model had adequate fit, which suggested that both the number and pattern of factors were equivalent across groups (see Table 3). Metric invariance was tested and the $\Delta\chi^2$ (6) = 3.098, $p = .796$ and ΔCFI = .000 values indicated that the constraint of factor loadings did not result in a significantly worse model fit compared with configural invariance, which supported metric invariance. Subsequently, scalar invariance was examined and the difference in the model fit between scalar invariance and the configural model was significant: $\Delta\chi^2$ (15) = 61.435, $p < .001$ and ΔCFI = −.014, indicating that scalar invariance was not achieved. A subsequent analysis was performed to determine

Table 3. Measurement invariance analysis.

	CFI	SRMR	RMSEA (90% CI)	χ^2	*df*	Δ*df*	$\Delta\chi^2$	ΔCFI
Configural	.975	.035	.056 (.045–.067)	129.151	42			
Metric	.975	.036	.051 (.041–.062)	132.249	48	6	3.098, $p = .796$	.000
Scalar	.961	.037	.059 (.050–.069)	190.585	57	15	61.435, $p < .001$	.014
Scalar (partial)	.972	.035	.051 (.042–.061)	148.595	54	12	19.445, $p = .078$	.003

CFI = Comparative Fit Index. SRMR = Standardized Root Mean Square Residual. RMSEA (90% CI) = Root Mean Square Error of Approximation (90% confidence interval). χ^2 = chi-square. *df* = degrees of freedom. $\Delta\chi^2$, Δ*df*, and ΔCFI were the difference between each alternative and the configural model.

Table 4. Means and standard deviations for boys and girls.

	Boys	Girls			
	M (SD)	*M (SD)*	*t-test*	*p*	*d*
Scales					
Inhibit	12.17 (3.22)	11.61 (2.91)	2.378	.018	0.18
Self-Monitor	6.32 (1.96)	5.94 (1.66)	2.791	.005	0.21
Shift	11.74 (2.78)	11.64 (2.58)	0.495	.621	0.03
Emotional Control	11.64 (3.26)	11.60 (3.03)	0.174	.862	0.01
Initiate	7.76 (2.15)	7.32 (1.95)	2.806	.005	0.21
Working Memory	12.28 (3.47)	11.93 (3.26)	1.388	.166	0.10
Plan/Organize	13.18 (3.29)	12.58 (3.15)	2.445	.015	0.18
Task Monitor	8.90 (2.25)	7.90 (2.04)	6.132	<.001	0.46
Org. Materials	8.59 (2.35)	8.56 (2.12)	0.171	.864	0.01
Indexes					
BRI	18.47 (4.74)	17.55 (4.16)	2.702	.007	0.20
ERI	23.38 (5.39)	23.25 (4.95)	0.330	.741	0.02
CRI	50.78 (11.40)	48.28 (10.68)	2.949	.003	0.22
GEC	92.57 (18.71)	88.88 (17.22)	2.640	.008	0.20

Effect sizes: Cohen's *d*. BRI = Behavior Regulation Index. ERI = Emotion Regulation Index. CRI = Cognitive Regulation Index. GEC = Global Executive Composite.

which intercepts were non-invariant and it revealed three intercept parameters that were not operating equivalently across groups (Initiate, Plan/Organize, and Task Monitor scales). Invoking the partial-measurement invariance strategy (Byrne & van de Vijver, 2010; Vandenberg & Lance, 2000), the three non-invariant intercepts were allowed to be freely estimated in each group and the partial scalar invariance was achieved: $\Delta\chi^2$ (12) = 19.445, $p = .078$ and ΔCFI = −.003.

Gender differences

After confirming the measurement invariance (partial scalar invariance) across gender, we examined the differences between boys and girls on the scales and indexes scores of the BRIEF2 Parent Form (see Table 4). Independent samples *t-test* revealed significant group differences (boys > girls) in five scales (Inhibit, Self-Monitor, Initiate, Plan/Organize, and Task Monitor), two indexes (BRI and CRI), and GEC.

Discussion

The BRIEF2 is a rating scale that evaluates everyday behaviors associated with EF in children and adolescents (aged 5–18 years) in home and educational environments. The purpose of this study was to investigate the factor structure and the measurement invariance across gender of the BRIEF2 Parent Form in Portuguese TDC.

Although the BRIEF and BRIEF2 have been the most popular instruments for measuring everyday EF in clinical, psychoeducational, and research settings, factor analysis studies have found different factor structures (e.g., Jiménez & Lucas-Molina, 2019; Lace et al., 2021; Pérez-Salas et al., 2016; Usher et al., 2016). Thus, the first main objective of the present study was to examine the dimensional structure of BRIEF2 in healthy Portuguese children. The original factor structure with three factors and nine scales was confirmed in standardization studies (Gioia et al., 2015, 2017), clinical samples

(Jiménez & Lucas-Molina, 2019; Parhoon et al., 2021; Shum et al., 2021), and TDC (Muñoz & Filippetti, 2021). For example, Jiménez and Lucas-Molina (2019) found that the three-factor model with nine scales showed better fit than the two- and one-factor models with eight or nine scales in at-risk children from the Dominican Republic. Muñoz and Filippetti (2021) also observed, in a sample of Chilean TDC, that the three-factor model best fit the data in comparison with two- and one-factor models.

In contrast, Lace et al. (2021) found that a bifactor two-factor structure provided the best fit in a clinical sample of children from the United States, whereas Cumming et al. (2021) found little evidence for the original three-correlated-factor model in children at risk for developing emotional and behavioral disorders. Interestingly, Jacobson et al. (2020) through exploratory factor analysis and CFA in clinically referred children found that the BRIEF2 is better operationalized with six (or fewer) scales rather than the original nine scales.

A preliminary analysis of the modification indices suggested the addition of three error covariances (Inhibit – Emotional Control, Inhibit – Task Monitor, and Shift – Initiate), and these data make empirical and statistical sense because these scales may share some underlying executive processes and are strongly correlated. We found large correlation coefficients between Inhibit and Emotional Control ($r = .645$), Inhibit and Task Monitor ($r = .523$), Shift and Initiate ($r = .573$). Interestingly, Gioia et al. (2015) also added three error covariances (Inhibit – Emotional Control, Inhibit – Working Memory, and Inhibit – Organization of Materials) in the factor structure of the standardization sample of BRIEF2 Parent and Teacher Forms.

We examined five competing factor models through CFA and the three-factor models (correlated and hierarchical) were the most interpretable and parsimonious factor solutions. In addition, we also found evidence of convergent validity based on the AVE values (≥.58), and the correlations coefficients between the latent factors and between the scales of the same factor were adequate. Thus, our findings gave support for the original dimensional structure of the BRIEF2 in a sample of TDC.

Although the goodness-of-fit indices demonstrated an adequate fit, we found a slight elevation of the RMSEA value (.078), which is expected given the results presented in several studies with BRIEF and BRIEF2 (e.g., .070 to .130 in Gioia et al., 2015: p. 082 in Jiménez & Lucas-Molina, 2019: p. 110 in Lace et al., 2021: p. 100 in Parhoon et al., 2021: p. 143 in Pérez-Salas et al., 2016). A possible explanation for the large RMSEA values may be the presence of strong correlations between the scales from different factors (Jiménez & Lucas-Molina, 2019; Sun et al., 2018).

We also found acceptable internal consistency values (range .73 to .85 for the nine scales, and .86 to .95 for the three indexes and GEC), which are close to those obtained in the standardization studies of the BRIEF2 Parent Form in the United States (range .80 to .91 for the nine scales, and .90 to .97 for the three indexes and GEC; Gioia et al., 2015) and Spain (range: from .66 to .87 for the scales, and.from .86 to .95 for the three indexes and GEC; Gioia et al., 2017).

The second main objective of the present study was to examine whether the original three-correlated-factor model would be equivalent across gender. The results from the multiple-group analysis supported configural invariance, which suggests that the number and pattern of factors were equivalent across boys and girls. The full metric invariance was also established, which indicates that the strength of the relation between the observed variable (scales) and their latent factors was equivalent across groups. The scalar invariance was then assessed to evaluate whether children who have the same score on a latent factor would obtain the same score on the observed variable regardless of their group membership. Three scales (Initiate, Plan/Organize, and Task Monitor scales) were not operating equivalently across groups, suggesting that boys may obtain a significantly different score on these three scales relative to girls with an equal score on their associated index. The findings from the multiple-group analysis demonstrated the partial measurement equivalence of the BRIEF2 Parent Form.

As referred by some authors (e.g., Moura et al., 2018; Sideridis et al., 2015; Wicherts, 2016), it is important to include measurement invariance during the validation process of cognitive measures because most of the psychological research and clinical practice involves between-groups comparisons. Indeed, measurement invariance must be demonstrated, at least at the factor loading level (i.e., metric invariance) before interpreting mean differences between groups. Surprisingly, no studies were conducted to examine measurement invariance of the BRIEF2 Parent Form across gender in TDC, even though gender differences have been found in at-risk children (Cumming et al., 2021; Jiménez & Lucas-Molina, 2019) and the standardization of the BRIEF and BRIEF2 includes norms for boys and girls separately (Gioia et al., 2000, 2015). To our knowledge, only one study performed a measurement invariance across gender in a socially vulnerable sample of primary school-aged children from the Dominican Republic (Jiménez & Lucas-Molina, 2019). This study also found support for the measurement invariance of the three-factor model of the BRIEF2 Parent Form.

Subsequently, we examined gender differences with boys showing greater EF difficulties than girls on five scales (Inhibit, Self-Monitor, Initiate, Plan/Organize, and Task Monitor), two indexes (BRI and CRI), and GEC. Taken together, the results from the measurement invariance and the inferential analyses highlighted the need to develop gender-specific norms for boys and girls in the Portuguese version of the BRIEF2 Parent Form. These findings were partly consistent with those obtained by Jiménez and Lucas-Molina (2019) that found statistically significant gender differences (boys > girls) in Shift, Initiate, Task-Monitor, and Organization of Materials scales. The BRIEF and BRIEF2 standardization studies (Gioia et al., 2000, 2015) and studies with clinical samples (Cumming et al., 2021) also revealed significant gender differences (boys > girls). Similar results are found in other studies carried out with the BRIEF. For example, Fournet et al. (2015) observed that boys had greater EF difficulties in all scales, except for Organization of Materials and Monitor. Huizinga and Smidts (2011) found that boys revealed lower executive skills than girls on the Shift scale and all the scales of the MI (except the Organization of Materials), as well in MI and GEC. Thus, the literature has consistently found that boys revealed greater EF difficulties than girls in the BRIEF and BRIEF2 scales.

Notwithstanding the relevance of the present study, it had some limitations that should be addressed in future studies. First, factor analytic studies have found that age is another relevant variable in the BRIEF and BRIEF2 (Fournet et al., 2015). Subsequent studies on the Portuguese population should analyze the effect of age on the factor structure and their interaction with gender. Second, future research should examine the concurrent validity of the BRIEF2 with other EF rating scales (e.g., Childhood Executive Functioning Inventory and Barkley Deficits in Executive Functioning Scale –

Children and Adolescents) and with performance-based measures of EF (e.g., Stroop Color and Word Test, Trail Making Test, and Tower of London). Third, the measurement equivalence between the Parent Form and the Teacher Form needs to be investigated. Lastly, it would be particularly interesting to explore the equivalence of the BRIEF2 factor structure between TDC and children with neurodevelopmental, neurological, or behavioral disorders (e.g., ADHD, specific learning disorder, autism spectrum disorder, epilepsy, traumatic brain injury, oppositional defiant disorder).

In conclusion, this study provides evidence regarding the adequate psychometric properties of the BRIEF2, suggesting that the parent version is useful for assessing executive functioning based on reports of behaviors observed by parents in Portuguese children. Specifically, these findings support the original three-correlated-factor model and demonstrated the measurement equivalence across gender.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Octávio Moura http://orcid.org/0000-0002-5857-6200
Cristina P. Albuquerque http://orcid.org/0000-0001-9055-9673
Marcelino Pereira http://orcid.org/0000-0002-1468-2124
Sofia Major http://orcid.org/0000-0002-4643-2170
Ana Filipa Lopes http://orcid.org/0000-0003-1091-1900
Manuela Vilar http://orcid.org/0000-0001-5245-7000
Maria João Seabra-Santos http://orcid.org/0000-0001-5035-0816
Mário R. Simões http://orcid.org/0000-0002-1311-1338

References

Baron, I. S. (2018). *Neuropsychological evaluation of the child: Domains, methods, & case studies* (2nd ed.). Oxford University Press.

Bednarz, H. M., Trapani, J. A., & Kana, R. K. (2020). Metacognition and behavioral regulation predict distinct aspects of social functioning in autism spectrum disorder. *Child Neuropsychology*, *26*(7), 953–981. https://doi.org/10.1080/09297049.2020.1745166

Blair, C., & Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. *Annual Review of Psychology*, *66*(1), 711–731. https://doi.org/10.1146/annurev-psych-010814-015221

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge.

Byrne, B. M., & van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing*, *10*(2), 107–132. https://doi.org/10.1080/15305051003637306

Chan, R. C. K., Shum, D., Toulopoulou, T., & Chen, E. Y. H. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*(2), 201–216. https://doi.org/10.1016/j.acn.2007.08.010

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Academic Press.

Cumming, M. M., Poling, D. V., Qiu, Y., Pham, A. V., Daunic, A. P., Corbett, N., & Smith, S. W. (2021). A validation study of the BRIEF2 among kindergarteners and first graders at-risk for behavior problems. *Assessment*, Advance online publication. https://doi.org/10.1177/10731911211032289

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29. https://doi.org/10.1037/1082-989X.1.1.16

Donders, J., DenBraber, D., & Vos, L. (2010). Construct and criterion validity of the Behaviour Rating Inventory of Executive Function (BRIEF) in children referred for neuropsychological assessment after paediatric traumatic brain injury. *Journal of Neuropsychology*, *4*(2), 197–209. https://doi.org/10.1348/174866409X478970

Egeland, J., & 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. https://doi.org/10.1080/09297041003601462

Fournet, N., Roulin, J.-L., Monnier, C., Atzeni, T., Cosnefroy, O., Le Gall, D., & 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. https://doi.org/10.1080/09297049.2014.906569

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *BRIEF: Behavior Rating Inventory of Executive Function* (1st ed.). PAR.

Gioia, G. A., Isquith, P. K., Retzlaff, P. D., & 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. https://doi.org/10.1076/chin.8.4.249.13513

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). *BRIEF-2: Behavior Rating Inventory of Executive Function* (2nd ed.). PAR.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2017). BRIEF-2. Evaluación conductual de *la Función ejecutiva. Manual técnico (M. J. Maldonado, M. C. Fournier, R. Martínez-Arias, J. González-Marquéz, J. M. Espejo-Saavedra, Y. P. Santamaría)*. TEA Ediciones.

Granader, Y., Wallace, G. L., Hardy, K. K., Yerys, B. E., Lawson, R. A., Rosenthal, M., Wills, M. C., Dixon, E., Pandey, J., Penna, R., Schultz, R. T., & Kenworthy, L. (2014). Characterizing the factor structure of parent reported executive function in autism spectrum disorders: The impact of cognitive inflexibility. *Journal of Autism and Developmental Disorders*, *44*(12), 3056–3062. https://doi.org/10.1007/s10803-014-2169-8

Guy, S. C., Gioia, G. A., & Isquith, P. K. (2004). *Behavior rating inventory of executive function: Self-report version*. Psychological Assessment Resources.

Halvorsen, M., Mathiassen, B., Amundsen, T., Ellingsen, J., Brøndbo, P. H., Sundby, J., Steinsvik, O. O., & Martinussen, M. (2019). Confirmatory factor analysis of the behavior rating inventory of executive function in a neuro-pediatric sample and its application to mental disorders. *Child Neuropsychology*, *25*(5), 599–616. https://doi.org/10.1080/09297049.2018.1508564

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huizinga, M., & 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. https://doi.org/10.1080/09297049.2010.509715

Jacobson, L. A., Kalb, L. G., & Mahone, E. M. (2020). When theory met data: Factor structure of the BRIEF2 in a clinical sample. *The Clinical Neuropsychologist*, *34*(1), 243–258. https://doi.org/10.1080/13854046.2019.1571634

Jiménez, A., & Lucas-Molina, B. (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. https://doi.org/10.1080/09297049.2018.1512962

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*(3), 213–233. https://doi.org/10.1007/s11065-007-9040-z

Lace, J. W., Seitz, D. J., Austin, T. A., Kennedy, E. E., Ferguson, B. J., & Mohrland, M. D. (2021). The dimensionality of the behavior rating inventory of executive function, second edition in a clinical sample. *Applied Neuropsychology: Child*, Advance online publication. https://doi.org/10.1080/21622965.2021.1910950

Lewis, M. P., Simons, G. F., & Fennig, C. D. (2015). *Ethnologue: Languages of Africa and Europe* (18th ed.). SIL International.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Masten, A. S., Herbers, J. E., Desjardins, C. D., Cutuli, J. J., McCormick, C. M., Sapienza, J. K., Long, J. D., & Zelazo, P. D. (2012). Executive function skills and school success in young children experiencing homelessness. *Educational Researcher*, *41*(9), 375–384. https://doi.org/10.3102/0013189X12459883

Moura, O., Simões, M. R., & Pereira, M. (2014). Executive functioning in children with developmental dyslexia. *The Clinical Neuropsychologist*, *28*(S1), 20–41. https://doi.org/10.1080/13854046.2014.964326

Moura, O., Albuquerque, C. P., Pinho, M. S., Vilar, M., Lopes, A. F., Alberto, I., Pereira, M., Santos, M. J. S., & Simões, M. R. (2018). Factor structure and measurement invariance of the Coimbra Neuropsychological Assessment Battery (BANC). *Archives of Clinical Neuropsychology*, *33*(1), 66–78. https://doi.org/10.1093/arclin/acx052

Muñoz, M. P., & Filippetti, V. A. (2021). Confirmatory factor analysis of the BRIEF-2 parent and teacher form: Relationship to performance-based measures of executive functions and academic achievement. *Applied Neuropsychology: Child*, *10*(3), 219–233. https://doi.org/10.1080/21622965.2019.1660984

Parhoon, K., Moradi, A., Alizadeh, H., Parhoon, H., Sadaphal, D. P., & Coolidge, F. L. (2021). Psychometric properties of the Behavior Rating Inventory of Executive Function, second edition (BRIEF2) in a sample of children with ADHD in Iran. *Child Neuropsychology*, *28*(4), 427–436. Advance online publication. https://doi.org/10.1080/09297049.2021.1975669

Pérez-Salas, C. P., Ramos, C., Oliva, K., & 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. https://doi.org/10.1177/0031512516650441

Pordata. (2018). *População residente segundo os censos: Total e por dimensão dos lugares*. Fundação Francisco Manuel dos Santos.

Riggs, N. R., Blair, C. B., & Greenberg, M. T. (2004). Concurrent and 2-year longitudinal relations between executive function and the behavior of 1st and 2nd grade children. *Child Neuropsychology*, *9*(4), 267–276. https://doi.org/10.1076/chin.9.4.267.23513

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., & Howlin, P. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition*, *71*(3), 362–368. https://doi.org/10.1016/j.bandc.2009.06.007

Shum, K.-K.-M., Zheng, Q., Chak, G. S., Kei, K.-T.-L., Lam, C.-W.-C., Lam, I.-K.-Y., Lok, C. S. W., & Tang, J.-W.-Y. (2021). Dimensional structure of the BRIEF2 and its relations with ADHD symptoms and task performance on executive functions in Chinese children. *Child Neuropsychology*, *27*(2), 165–189. https://doi.org/10.1080/09297049.2020.1817355

Sideridis, G. D., Tsaousis, I., & Al-Harbi, K. A. (2015). Multi-Population invariance with dichotomous measures: Combining multi-group and MIMIC methodologies in evaluating the general aptitude test in the Arabic language. *Journal of Psychoeducational Assessment*, *33*(6), 568–584. https://doi.org/10.1177/0734282914567871

Slick, D. J., Lautzenhiser, A., Sherman, E. M. S., & 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. https://doi.org/10.1080/09297040600611320

Sun, X., Wu, Z., Cao, Q., Qian, Y., Liu, Y., Yang, B., Chang, S., Yang, L., & 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*, Article 7620. https://doi.org/10.1038/s41598-018-26042-y

Usher, A. M. L., Leon, S. C., Stanford, L. D., Holmbeck, G. N., & 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. https://doi.org/10.1080/09297049.2015.1060956

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

Wasserman, T., & Wasserman, L. D. (2013). Toward an integrated model of executive functioning in children. *Applied Neuropsychology: Child*, *2*(2), 88–96. https://doi.org/10.1080/21622965.2013.748394

Wicherts, J. M. (2016). The importance of measurement invariance in neurocognitive ability testing. *The Clinical Neuropsychologist*, *30*(7), 1006–1016. https://doi.org/10.1080/13854046.2016.1205136

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, *57*(11), 1336–1346. https://doi.org/10.1016/j.biopsych.2005.02.006

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*(4), 354–360. https://doi.org/10.1111/j.1750-8606.2012.00246.x

Zelazo, P. D. (2020). Executive function and psychopathology: A neurodevelopmental perspective. *Annual Review of Clinical Psychology*, *16*(1), 431–454. https://doi.org/10.1146/annurev-clinpsy-072319-024242