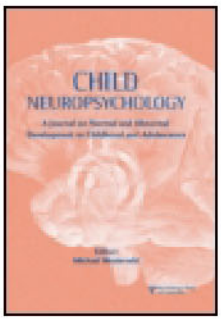

Child Neuropsychology

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ncny20

Brief Report: Validation of Catalan Version of BRIEF-P

Albert Bonillo [a], Eva Angelina Araujo Jiménez [a], María Claustre Jané Ballabriga [a], Connie Capdevila & Ramon Riera [b]

[a] Autonomous University of Barcelona, Bellaterra, Spain

[b] Equip d'Assessorament i Orientació Psicopedagògica (Educational and Psycho-pedagogical Team EAP), Osona, Vic, Spain

Available online: 03 Oct 2011

To cite this article: Albert Bonillo, Eva Angelina Araujo Jiménez, María Claustre Jané Ballabriga, Connie Capdevila & Ramon Riera (2011): Brief Report: Validation of Catalan Version of BRIEF-P, Child Neuropsychology, DOI:10.1080/09297049.2011.613808

To link to this article: http://dx.doi.org/10.1080/09297049.2011.613808

Child Neuropsychology, 2011, i*First*, 1–9
http://www.psypress.com/childneuropsych
ISSN: 0929-7049 print / 1744-4136 online
http://dx.doi.org/10.1080/09297049.2011.613808

Psychology Press
Taylor & Francis Group

Brief Report: Validation of Catalan Version of BRIEF-P

Albert Bonillo[1], Eva Angelina Araujo Jiménez[1], María Claustre Jané Ballabriga[1], Connie Capdevila, and Ramon Riera[2]

[1]Autonomous University of Barcelona, Bellaterra, Spain
[2]Equip d'Assessorament i Orientació Psicopedagògica (Educational and Psychopedagogical Team EAP), Osona, Vic, Spain

The BRIEF-P is one of the most used instruments to measure Executive Function (EF). This report is aimed at showing the psychometric properties of the Catalan version of it. To do this, a random cluster sampling was carried out in Osona (Catalonia), recruiting 417 teachers and 408 parents of children aged 3 to 6 years. As with the original instrument, results show that reliability is excellent and that girls have a better EF than boys in the following areas: Working Memory, Plan/Organize, and Inhibit. Since in some areas Catalonia's preschool children show EF values slightly better than those of the reference sample, a new scale for using it in this population is provided.

Keywords: Executive function; Preschool children; BRIEF-P; Psychometric validation.

INTRODUCTION

Executive Function (EF) is a set of neurocognitive processes mediated by the prefrontal cortex to guide behavior towards the achievement of goals (Aarnoudse-Moens, Smidts, Oosterlaan, Duivenvoorden, & Weisglas-Kuperus, 2009; Welsh, 2002, as cited in Wåhlstedt, Thorell, & Bohlin, 2008). EF includes focusing attention, recognition of priorities, formulation of goals, planning activities, implementation of plans, self-regulation, inhibitory control, flexibility, and self-evaluation of the results of action taken (P. Anderson, 2002; Senn, Espy, & Kaufmann, 2004). EF begins to develop at an early age. At 3 years of age, executive activities can be seen, for instance, rule following, highly complex cognitive activities, control of emotions, and problem solving (Espy, 2004; Zelazo, 2004).

In children, EF deficiencies are associated with attention deficit disorder with hyperactivity (Brocki, Eninger, Thorell, & Bohlin, 2009), oppositional defiant disorder (van Goozen et al., 2004), antisocial behavior (Wåhlstedt et al., 2008), autism,

This work had the collaboration of the "Equip d'Assessorament i orientació Psicopedagògica d'Osona, EAP," and it was funded by the "Fondos de Investigación Sanitario FIS" number 070027.

Thanks to Eva Penello Werner for her help in reviewing the CFA.

Address correspondence to Albert Bonillo, Departament de Psicolobiologia i Metodologia de les Ciències del Comportament, Universitat Autònoma de Barcelona, Edifici B Facultat Psicologia, Campus UAB, 08193 Bellaterra (Cerdanyola del Vallès), Spain. E-mail: Albert.Bonillo@uab.cat

Asperger syndrome, Tourette syndrome, and obsessive-compulsive disorder, among others (Pennington & Ozonoff, 1996).

Although the theoretical construct of EF is a cognitive psychology classic, there are not many standardized instruments measuring it. Those measuring it are rather recent (Huizinga & Smidts, 2011). Among them, the most known and used is probably the Behavior Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000). It has versions for preschoolers (3–6 years of age), children (5–18 years of age), and adults (18–90 years of age).

Possibly because of the recency of these instruments, there are very few adaptations to languages other than English. The children's version has been translated into Dutch (Huizinga & Smidts, 2011) and Hebrew (Linder, Kroyzer, Maeir, Wertman-Elad, & Pollak, 2010). Nevertheless, in the case of the preschool version presented in this work, we do not know of any other translation whose psychometric data have been published.

The aim of this brief report is to analyze the psychometric properties of the Catalan version of the BRIEF-P (Gioia, Espy, & Isquith, 2003) and to provide a validated scale that clinicians can use in their daily practice. It should be noted that the territory where people speak Catalan has 13 million inhabitants, of which 11 million speak it and for many of them Catalan is their usual language. Due to the number of speakers, Catalan is the ninth most spoken language in the EU, ahead of others such as Portuguese and Swedish (Eurostat, 2010).

METHOD

We carried out a random cluster sampling in Osona (Catalonia, NE of Spain). We recruited approximately 10% of the population between 3 and 6 years of age. The main objective was to obtain a representative sample of 400 children. With this *N*, it is possible to detect differences between means of 0.05 when comparing scales, given a population of 0.4 σ (maximum appearing in Isquith, Crawford, Espy, & Gioia, 2005) with a power (1-β) greater than .8 and a first kind error of $\alpha = .05$ (Lemeshow, Hosmer, Klar, & Lwanga, 1990).

Taking an average of 30 children per school, 13 Osona schools were randomly selected. Five of them refused to participate and were replaced using the same procedure. All parents and teachers of 444 children between 3 and 6 years of age enrolled in these schools agreed to participate. After receiving the BRIEF Preschool Version (BRIEF-P) to respond to in a self-administered way, 408 parents (92%) and 417 teachers (94%) returned it, and if they had any question it was answered by a member of our research team, with values below the critical validity scales: negativity (100% parents; 96.6% teachers) and inconsistency (95.5% fathers; 96.8% teachers). Of the 417 children whose teachers participated, 204 were girls (48.9%) and 213 were boys (51.1%); 143 (34.3%) were in P3 level (3–4 years of age), 132 (31.7%) in P4 level (4–5 years of age), and 142 (34.1%) in P5 level (5–6 years of age).

The Behavior Rating Inventory of Executive Function-Preschool Version (BRIEF-P; Gioia et al., 2003) comprises a questionnaire for parents and one for teachers. Each questionnaire has 63 items with a 3-point scale (1 = Never; 2 = Sometimes; 3 = Often) divided into five clinical scales (Inhibition, Shift, Emotional Control, Working Memory, and Plan/Organize), which form three broader indexes (Inhibitory self-control: groups' inhibition and emotional control; Flexibility: groups' shift and emotional control; Emergent Metacognition: groups' working memory and plan/organize), two validity scales (Inconsistency and Negativity), and a Global Execution combined score. High scores on

any scale are interpreted as an indicator of poor EF. These can be calculated based on the sum of items or the mean of items. We chose the latter option, since it allows for keeping the same metric at all scales, regardless of how many items they have.

After obtaining permission from the authors of the BRIEF-P to develop the Catalan translation, a clinical psychologist whose first language is Catalan, who is Anglophone and an expert in Executive Function, translated the items. A clinician, whose first language is English and also speaks Catalan, conducted the back-translation. The latter was sent to the authors for them to assess their equivalence with the original. The few corrections proposed by the authors were incorporated into the Catalan version.

All statistic data were analyzed with SPSS software for Windows, version 19.0, except for the calculation of the Confirmatory Factor Analysis (CFA), which was calculated with Mplus v.6 (Muthén & Muthén, 2010).

RESULTS AND DISCUSSION

Table 1 shows the statistics discussed in this section, unless otherwise specified.

First, as a measure of internal consistency of the first-order clinical scales, we calculated Cronbach's alpha (1951) based on the items for parents and teachers. Only 2 of the 18 calculated values were lower than .8 (Organize & Shift Plan). The remainders, in many cases, were higher than .9. This is consistent with excellent internal consistency, just as occurs in the original instrument (Gioia et al., 2003).

Second, *F* tests were applied (Shrout & Fleiss, 1979) to compare each of the calculated Cronbach's alpha with those published in Isquith, Gioia, and Espy (2004). There were no statistically significant differences. This leads us to conclude that not only the reliability is excellent, but also it is statistically equivalent to the original instrument. Since the studied version is a translation, this result provides evidence of the validity of our version and of the transculturality of the construct.

Third, a second-order factor analysis based on the five clinical scales for parents and teachers was carried out. As in Isquith, Gioia, and Espy (2004), the solution of three broader indexes was tested by using a Promax rotation, given the expected correlation between factors (Hendrickson & White, 1964). The value of r^2 obtained for both informants (89% parents, 93% teachers) was outstanding. This indicates that the factorial solution summarizes the original data exceptionally well. Consistent with this excellent correspondence, the factor loadings were high. In many cases, they were higher than .9, and in all cases they were always above .3. Traditionally, this value is considered critical. The lowest factor loadings were given in the weight of emotional control on the flexibility factor, in both parents and teachers, as in the previously cited article. Broader index correlations between the same informant and correlations between informants' factors were moderate, as in the article by Gioia et al. (2003). All this confirms that the constructs of the scales are the same for parents and teachers, but that both provide different information because the child is assessed in different contexts (Viñas et al., 2008).

Fourth, we try to perform a Confirmatory Factor Analysis (CFA) of the BRIEF-P's complete structure (i.e., 63 items, the five clinical scales and the three broader indexes) in order to calculate an overall index that measures the goodness of fit. However, this was not possible for data from parents and for data from teachers, since correlations between latent variables were consistently greater than 1. In the literature these cases are known as *Heywood cases* or *offending estimates* (Dillon, Kumar, & Mulani, 1987; Satorra, 1990) and they may occur in various settings: because of a bad model specification, due to an insufficient size sample, due to the high complexity of the model to be estimated, or due

Table 1 Descriptive, Reliability, and Factor Loadings for a Three-Factor Model: Parent and Teacher Ratings.

		Parents Sample ($n = 408$)						Teachers Sample ($n = 417$)					
			Factor loadings ($r^2 = 89\%$)			Mean (*SD*)			Factor loadings ($r^2 = 93\%$)			Mean (*SD*)	
Scale	*r* Parent-Teachers	Cronbach α	Factor 1	Factor 2	Factor 3	Boys	Girls	Cronbach α	Factor 1	Factor 2	Factor 3	Boys	Girls
Working Memory	.386	.878	.934			1.38 (0.31)	1.25 (0.26)*	.958	.969			1.37 (0.44)	1.17 (0.30)*
Plan/Organize	.268	.770	.917			1.48 (0.33)	1.37 (0.28)*	.909	.966			1.36 (0.41)	1.16 (0.28)*
Shift	.179	.720		.974		1.38 (0.30)	1.36 (0.28)	.830		.978		1.25 (0.30)	1.19 (0.26)
Emotional Control	.451	.803		.445	.942	1.49 (0.33)	1.48 (0.34)	.917		.623	.885	1.36 (0.44)	1.30 (0.37)
Inhibit	.296	.872			.766	1.60 (0.37)	1.43 (0.29)*	.945			.912	1.45 (0.49)	1.17 (0.23)*
Factor 1	.312	.908				1.43 (0.30)	1.31 (0.26)*	.969				1.36 (0.42)	1.16 (0.28)*
Factor 2	.189	.839	.208			1.43 (0.27)	1.42 (0.27)	.917	.271			1.30 (0.33)	1.24 (0.28)
Factor 3	.411	.897	.554	.320		1.54 (0.31)	1.45 (0.28)*	.953	.528	.413		1.40 (0.43)	1.23 (0.27)*
GEC	.330	.939				1.46 (0.26)	1.38 (0.22)*	.971				1.36 (0.34)	1.20 (0.22)*

Note. r^2 = % of variance explained by factorial structure of second order; r = Pearson Correlation; Factor 1 = Emergent Metacognition; Factor 2 = Flexibility; Factor 3 = Inhibitory Self-Control; GEC = Global Executive Composite (Total Score).

*Indicates statistical differences between gender.

to a combination of several of the above mentioned causes (Anderson & Gerbing, 1984). Given the success of the second-order factor analysis, we believe that the sample size (about 420 cases to estimate 198 parameters when it is recommended 5 to 10 cases per parameter; Bentler & Chou, 1987) along with the complexity of the model to be estimated (the three broader indexes were calculated based on five clinical scales) may explain this effect. It would have been interesting to see whether the same thing happened in normative samples, but it was not possible to calculate it because the intercorrelations of the 63 items were not published. With them, it would have been possible to calculate the models without having individualized data.

Fifth, as a consequence of the above, we calculated a one-dimensional CFA, where the combined Global Execution score is calculated from the intercorrelations of the five clinical scales. This was done for data from parents and teachers and for normative samples, as these intercorrelations are published (see Gioia et al., 2003, pp. 55–56). Model fit was assessed with several fit indices: comparative fit index (CFI; Bentler, 1990), non-normed fit index (TLI or NNFI; Bentler & Bonett, 1980), standardized root-mean-square residual (SRMR; Bentler, 1990), and root-mean-square error of approximation (RMSEA; Steiger, 1990). For the RMSEA, values of .05 or lower indicate excellent fit and between .05 and.08 adequate fit (Browne & Cudeck, 1993). Hu and Bentler (1999) suggested that CFI-NNFI values of .95 or higher indicate good fit; though other researchers have suggested cutoffs of .08 for SRMR-RMSEA and .90 for CFI-NNFI (e.g., Jöreskog, Sörbom, du Toit, & du Toit, 2000; Marsh, Hau, & Wen, 2004).

In the parent sample, the fit index for the one-factor CFA model at the clinical scale level are: CFI = .903; NNFI = .806; RMSEA = .210; SRMR = .065; in the normative sample they are: CFI = .887; NNFI = .773; RMSEA = .214; SRMR = .070. In the teacher sample, they are: CFI = .851; NNFI = .702; RMSEA = .313; SRMR = .098; in the normative sample they are: CFI = .840; NNFI = .679; RMSEA = .303; SRMR = .107. From these data, we can say that the models of our samples have very similar fit indexes; although they are slightly better than those of normative samples. However, neither one case nor the other is good. From these data, we can say that the models of our samples have very similar fit indexes; although they are slightly better than those of normative samples, in one case or another they are not optimal, especially regarding RMSEA and, at a lesser extent, SRMR.

We must emphasize that these fit indexes should not be interpreted as an adjustment of the structure of the original model, since it is not possible to add the three broader indexes to the five clinical scales because it would be an overidentified model. However, this strategy does allow us to assess the fit of the global score and the clinical scales in the Catalan version with respect to the original, and the main objective of this study is to assess whether the reliability and validity of both are equivalent. Now well, we believe that it is clear that the structure of first and second order of the BRIEF-P should be investigated more deeply, which is not required for the items since it seems already clear. There is no published article where the overall fit of the BRIEF-P is calculated with a CFA. Even for other versions of the brief, the papers are scarce.

There are no articles in which such a fit is calculated. To the best of our knowledge, there are only three publications doing it, and all of them are child versions: One is in Chinese (Qian & Wang, 2009), another uses a clinical sample (Gioia, Isquith, Retzlaff, & Espy, 2002), and the last one uses a mixed clinical-community sample and applies the Norwegian translation (Egeland & Øyvind, 2010). Clearly, this aspect should be deepened in order to ensure that the structure of the BRIEF corresponds to what is being measured.

Sixth, to assess differences in scores attributable to gender, age or their interaction, it was carried out a factor analysis of variance (GLM) of the clinical scales, the broader indexes, and of the total score of the respondents, using Bonferroni correction for avoiding to increase the first kind error (Dunn, 1961). Neither interaction nor age were significant in any case.

As for gender, differences were significant but practically irrelevant in the clinical scales Working Memory, Plan/Organize, and Inhibit, both for parents and teachers. The calculated statistics for parents were Working Memory, $F(1, 353) = 14.29$, $p < .0005$, $\eta^2 = .04$; Plan/Organize, $F(1, 145) = 7.37$, $p < .007$, $\eta^2 = .02$, and Inhibit, $F(1, 353) = 20.34$, $p < .0005$, $\eta^2 = .05$. The calculated statistics for teachers were Working Memory, $F(1, 341) = 24.58$, $p < .0005$, $\eta^2 = .07$; Plan/Organize, $F(1, 353) = 21.34$, $p < .0005$, $\eta^2 = .06$, and Inhibit, $F(1, 353) = 34.37$, $p < .0005$, $\eta^2 = .09$. Isquith, Gioia, and Espy (2004) also found that teachers reported poorer executive function in boys than girls in these three areas, whereas, when parents reported, only inhibition was worse in boys. In our case, we found that Working Memory and Plan/Organize also were worse in boys than in girls. This may be because our sample size was slightly higher and, consequently, the tests had more power.

Seventh, to assess to what extent the means of our sample were statistically equivalent to each of the dimensions of the article of Isquith, Gioia, and Espy (2004), one-sample *t*-test (Bland, 2000) for each dimension, informant, and gender was carried out using again the Bonferroni correction. When parents valued boys, there was a significant difference in the size of Shift. When it came to parents and girls, there were significant differences in all means except for Inhibit. When teachers valued boys, only the differences between Shift and Plan/Organize were significant. If it came to girls, all means were different. In all these cases, the means of our sample were lower than those of the reference. This might indicate better executive function, but the difference is clinically irrelevant and largely explained by the statistical power given by the sample size.

Given these differences, and although the Catalan's clinicians can obtain the critical values of their population for each scale from Table 1 (equivalent to a T score > 65, i.e., adding 1.5 standard deviation to the mean), Table 2 shows the raw scores that correspond to the most commonly used percentiles (by gender) for the Catalan population, which were calculated from our sample.

One of the major contributions of this work is that the EF of the Catalan-speaking preschool children can be assessed in their native language, which, in many cases, is also the native language of their teachers. All the inhabitants of Catalonia can speak Spanish and most understand Catalan, although not all speak Catalan. Now, half the population has Catalan as native language, a third of the population has Catalan as exclusive language, and the education system is taught entirely in Catalan. A similar situation, but with less diffusion of Catalan, occurs in other regions of Spain such as Valencia and the Balearic Islands, and in a lesser extent in parts of northeastern France, and in Alghero, Sardinia, Italy (Idescat, 2010).

Once assured the reliability of the translation, our next step should be to ensure its validity. This will require recruiting new community and clinical samples, and measuring all the variables that are usually related to EF. Given the close relationship between EF and various disorders, adjusting the data is the only way to avoid the confusion inherent in any statistical nonexperimental study.

Original manuscript received March 23, 2011
Revised manuscript accepted July 17, 2011

Table 2 Normative Values by Percentile, Informant, and Gender of Preschooler: Catalan Sample.

		Boys							Girls						
		25	50	75	85	90	95	99	25	50	75	85	90	95	99
Parents	Inhibit	1.31	1.56	1.75	1.92	2.01	2.22	2.70	1.19	1.44	1.59	1.69	1.81	1.94	2.25
	Shift	1.10	1.30	1.50	1.60	1.70	1.90	2.31	1.10	1.30	1.50	1.60	1.70	1.90	2.30
	Emotional Control	1.20	1.40	1.70	1.80	1.90	2.00	2.51	1.20	1.40	1.70	1.90	1.90	2.09	2.40
	Working Memory	1.12	1.29	1.53	1.71	1.82	1.94	2.15	1.06	1.18	1.35	1.53	1.59	1.82	2.06
	Plan/Organize	1.20	1.40	1.60	1.80	1.90	2.10	2.51	1.20	1.30	1.60	1.70	1.80	2.00	2.10
	F1	1.31	1.51	1.72	1.84	1.95	2.18	2.47	1.26	1.44	1.65	1.71	1.76	1.97	2.25
	F2	1.25	1.42	1.55	1.66	1.80	1.95	2.30	1.21	1.40	1.55	1.70	1.75	1.90	2.35
	F3	1.20	1.38	1.59	1.73	1.86	1.99	2.31	1.11	1.24	1.47	1.61	1.69	1.81	2.08
	GEC	1.27	1.44	1.60	1.72	1.83	1.95	2.16	1.20	1.36	1.53	1.60	1.66	1.81	1.96
Teachers	Inhibit	1.06	1.44	1.59	1.69	1.81	1.94	2.25	1.00	1.06	1.25	1.35	1.44	1.63	2.00
	Shift	1.00	1.30	1.50	1.60	1.70	1.90	2.30	1.00	1.10	1.30	1.40	1.60	1.80	1.92
	Emotional Control	1.00	1.40	1.70	1.90	1.90	2.09	2.40	1.00	1.10	1.50	1.70	1.80	1.80	2.41
	Working Memory	1.00	1.18	1.35	1.53	1.59	1.82	2.06	1.00	1.00	1.18	1.29	1.47	1.84	2.00
	Plan/Organize	1.00	1.30	1.60	1.70	1.80	2.00	2.10	1.00	1.00	1.20	1.30	1.44	1.80	2.00
	F1	1.06	1.44	1.65	1.71	1.76	1.97	2.25	1.00	1.14	1.38	1.51	1.65	1.85	2.19
	F2	1.05	1.40	1.55	1.70	1.75	1.90	2.35	1.00	1.10	1.40	1.56	1.70	1.84	2.10
	F3	1.03	1.24	1.47	1.61	1.69	1.81	2.08	1.00	1.03	1.19	1.39	1.62	1.85	2.10
	GEC	1.09	1.36	1.53	1.60	1.66	1.81	1.96	1.02	1.12	1.30	1.43	1.52	1.68	2.08

Note. Factor 1 = Emergent Metacognition; Factor 2 = Flexibility; Factor 3 = Inhibitory Self-Control; GEC = Global Executive Composite (Total Score).

REFERENCES

Aarnoudse-Moens, C. S. H., Smidts, D. P., Oosterlaan, J., Duivenvoorden, H. J., & Weisglas-Kuperus, N. (2009). Executive function in very preterm children at early school age. *Journal of Abnormal Child Psychology*, *37*(7), 981–993. doi:10.1007/s10802-009-9327-z

Anderson, J. C., & Gerbing, D. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, *49*, 155–173. doi:10.1007/BF02294170

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71–82. doi:10.1076/chin.8.2.71.8724

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. doi: 10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. doi: 10.1037/0033-2909.88.3.588

Bentler, P. M., & Chou, C. (1987). Practical issues in structural modeling. *Sociological Methods and Research*, *16*, 78–117. doi: 10.1037/0033-2909.88.3.588

Bland, M. (2000). *An introduction to medical statistics*. Oxford, UK: Oxford University Press.

Brocki, K. C., Eninger, L., Thorell, L. B., & Bohlin, G. (2009). Interrelations between executive function and symptoms of hyperactivity/impulsivity and inattention in preschoolers: A two year longitudinal study. *Journal of Abnormal Child Psychology*, *38*(2), 163–171. doi:10.1007/s10802-009-9354-9

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Alternative ways of assessing model fit* (pp. 136–162). Beverly Hills, CA: Sage.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. doi:10.1007/BF02310555

Dillon, W. R., Kumar, A., & Mulani, N. (1987). Offending estimates in covariance structure analysis: Comments on the causes and solutions to Heywood cases. *Psychological Bulletin*, *101*, 126–135. doi: 10.1037/0033-2909.101.1.126

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, *56*, 52–64.

Egeland, J., & Øyvind, F. (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), 249–257. doi: 10.1080/09297041003601462

Espy, K. A. (2004). Using developmental, cognitive, and neuroscience approaches to understand executive control in young children. *Developmental Neuropsychology*, *26*, 379–384. doi: 10.1207/s15326942dn2601_1

Eurostat. (2010). *Europe in figures: Eurostat YearBook 2010*. Brussels, Belgium: Author.

Gioia, G. A., Espy, K. A., & Isquith, P. K. (2003). *Behavior Rating Inventory of Executive Function-Preschool Version: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function BRIEF: Professional manual*. Lutz, FL: Psychological Assessment Resources.

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), 326–337.

Hendrickson, A. E., & White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *The British Journal of Statistical Psychology*, *17*, 65–70.

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. doi: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*, 51–66. doi: 10.1080/09297049.2010.509715

Idescat. (2010). *Anuari estadístic de Catalunya*. Barcelona, Spain: Author.

Isquith, P. K., Crawford, J. S., Espy, K. A., & Gioia, G. A. (2005). Assessment of executive function in preschool-aged children. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*(3), 209–215. doi:10.1002/mrdd.20075

Isquith, P. K., Gioia, G. A., & Espy, K. A. (2004). Executive Function in preschool children: Examination through everyday behavior. *Developmental Neuropsychology*, *26*(1), 403–422. doi: 10.1207/s15326942dn2601_3

Jöreskog, K. G., Sörbom, D., du Toit, S. H. C., & du Toit, M. (2000). *M. LISREL 8: New statistical features*. Lincolnwood, IL: Scientific Software International.

Lemeshow, St., Hosmer, D. W., Klar, J., & Lwanga, St. K. (1990). *Adequacy of sample size in health studies*. New York, NY: J. Wiley & Sons.

Linder, N., Kroyzer, N., Maeir, A., Wertman-Elad, R., & Pollak, Y. (2010). Do ADHD and executive dysfunctions, measured by the Hebrew version of Behavioral Rating Inventory of Executive Functions (BRIEF), completely overlap? *Child Neuropsychology*, *16*(5), 494–502. doi:10.1080/09297041003781884

Marsh, H. W., Hau, K. T., & 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 Modelling*, *11*, 320–341.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus User's Guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child and Psychiatry*, *37*, 51–87. doi: 10.1111/j.1469-7610.1996.tb01380.x

Qian, Y., & Wang, Y. (2009). Reliability and validity of the Behavior Rating Inventory of Executive Function Teacher Form for school-age children in China. *Chinese Mental Health Journal*, *23*(10), 742–747.

Satorra, A. (1990). Robustness issues in structural equation modeling: A review of recent developments. *Quality and Quantity*, *24*, 367–386.

Senn, T. E., Espy, K. A., & Kaufmann, P. M. (2004). Using path analysis to understand executive function organization in preschool children. *Developmental Neuropsychology*, *26*(1), 445–464. doi:10.1207/s15326942dn2601_5

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *2*, 420–428. doi: 10.1037/0033-2909.86.2.420

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*, 173–180.

van Goozen, S. H., Cohen-Kettenis, P. T., Snoek, H., Matthys, W., Swaab-Barneveld, H., & van Engeland, H. (2004). Executive functioning in children: A comparison of hospitalised ODD and ODD/ADHD children and normal controls. *Journal of Child Psychology and Psychiatry*, *45*(2), 284–292. doi:10.1111/j.1469-7610.2004.00220.x

Viñas, F., Jané, M. C., Canals, J., Esparó, G., Ballespí, S., & Domènech-Llaberia, E. (2008). Evaluación de la psicopatología del preescolar mediante el Early Childhood Invetory-4 (ECI-4): Concordancia entre padres y maestros. *Psicothema*, *20*, 481–486. Retrieved from http://dialnet.unirioja.es/servlet/articulo?codigo=2684842 [Assessment of psychopathology from preschool through the Early Childhood Invetory-4 (ECI-4): Concordance between parents and teachers]

Wåhlstedt, C., Thorell, L. B., & Bohlin, G. (2008). ADHD symptoms and executive function impairment: Early predictors of later behavioral problems. *Developmental Neuropsychology*, *33*(2), 160–178. doi:10.1080/87565640701884253

Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences*, *8*(1), 12–17. doi:10.1016/j.tics.2003.11.001