

Running head: Reliability and validity of BRIEF-P

Postprint de: Ezpeleta, L., Granero, R., Penelo, E., de la Osa, N., & Domènech, J. M. (2015). Behavior Rating Inventory of Executive Functioning–Preschool (BRIEF-P) Applied to Teachers: Psychometric Properties and Usefulness for Disruptive Disorders in 3-Year-Old Preschoolers. *Journal of Attention Disorders*, 19(6), 476–488.

<https://doi.org/10.1177/1087054712466439>

**Behavior Rating Inventory of Executive Functioning-Preschool (BRIEF-P)**

**Applied to Teacher: Psychometric Properties and Usefulness for Disruptive**

**Disorders in 3-year-old preschoolers**

Lourdes Ezpeleta<sup>1,2</sup>

Roser Granero<sup>1,3</sup>

Eva Penelo<sup>3</sup>

Núria de la Osa<sup>1,2</sup>

Josep M. Domènech<sup>1,3</sup>

<sup>1</sup>Unitat d'Epidemiologia i de Diagnòstic en Psicopatologia del Desenvolupament

<sup>2</sup>Departament de Psicologia Clínica i de la Salut

<sup>3</sup>Laboratori d'Estadística Aplicada, Departament de Psicobiologia i Metodologia de

les Ciències de la Salut

Universitat Autònoma de Barcelona, Barcelona, Spain

✉ Lourdes Ezpeleta

Departament de Psicologia Clínica i de la Salut. Edifici B

Universitat Autònoma de Barcelona

08193 Bellaterra (Barcelona), SPAIN

Phone: (34) 93 581 2883

E-mail: Lourdes.Ezpeleta@uab.cat

### **Acknowledgements**

Funding was from Spanish Ministry of Science and Innovation grant PSI2009-07542.

We thank the participating schools and families.

**Lourdes Ezpeleta, Ph.D.** is professor of child and adolescent psychopathology at the Universitat Autònoma de Barcelona (Spain). She is also the Director of the Unit of Epidemiology and Diagnostic in Developmental Psychopathology. Her field of research has been focused on developing and testing diagnostic instruments for epidemiological studies in child and adolescent psychopathology as well as on the study of risk factors of child and adolescent psychopathology.

**Roser Granero, Ph.D.** is an associate professor of methodology of behavioral sciences at the Universitat Autònoma de Barcelona (Spain). Her field of research has been focused on design, statistics and methodology for scientific analysis in Health Sciences. She is researcher at the Unit of Epidemiology and Diagnostic in Developmental Psychopathology (UED).

**Eva Penelo, Ph.D.** is lecturer of the Unit of Methodology at the Universitat Autònoma de Barcelona (Spain). Her research interests include applied psychometrics, instrument validation and structural equation modeling.

**Núria de la Osa, Ph.D.** is assistant professor of Psychological Assessment at the Universitat Autònoma de Barcelona (Spain). She is member of the Unit of Epidemiology and Diagnostic in Developmental Psychopathology. Her research interests involve children and adolescents clinical assessment and prevention.

**Josep Maria Domènech, Ph.D.** is professor of methodology of behavioral sciences at the Universitat Autònoma de Barcelona (Spain). He is also the Director of the Laboratori d'Estadística Aplicada i de Modelització at the Universitat Autònoma de Barcelona and a researcher at the Unit of Epidemiology and Diagnostic in Developmental Psychopathology (UED). His field of research has been focused on statistical models and methodology in Health Sciences.

### **Abstract**

**Objective:** We provide validation data on the Behavior Rating Inventory of Executive Functioning - Preschool version (BRIEF-P) in preschool children. **Method:** Teachers of a community sample of 620 3-year-olds, who were followed up at age 4, responded to the BRIEF-P, and parents and children answered different psychological measures.

**Results:** Confirmatory Factor Analysis achieved adequate fit of the original structure (5-first-order-factor plus 3-second-order-factor model) after excluding 4 items. The derived dimensions obtained satisfactory internal consistency, moderate convergent validity with psychopathology and temperament, and good ability to discriminate between children with ADHD. BRIEF-P scales were not associated with a performance-based measure of attention. The teacher's BRIEF-P adds significant clinical information for the diagnosis of ADHD ( $\Delta R^2$  from 5.3 to 15.3) when used with other instruments for the assessment of psychopathology, functional impairment or performance-based attention.

**Conclusions:** The BRIEF-P may be useful in the identification of preschool children, specifically those with ADHD, who might have a dysfunction in executive functioning.

**Key words:** Confirmatory factor analysis; Behavior Rating Inventory of Executive Functioning - Preschool version; executive functions; incremental validity; preschool children; reliability; validity.

**Behavior Rating Inventory of Executive Functioning-Preschool (BRIEF-P)  
Applied to Teacher: Psychometric Properties and Usefulness for Disruptive  
Disorders in 3-year-old preschoolers**

The Behavior Rating Inventory of Executive Functioning-Preschool Version (BRIEF-P) is an ecological measure for the assessment of executive functions in 2 to 5-year-old children as perceived by their parents or their teachers (Gioia, Espy, & Isquith, 2003; Gioia, Isquith, Guy, & Kenworthy, 2000). Together with the other versions of the questionnaire (BRIEF for parents/teachers of 5 to 18-year-olds, BRIEF-SR for 11 to 18-year-olds, and the BRIEF-Adult for ages 18 to 90), it permits the study, from a developmental perspective, of executive functioning as observed in daily functioning. The characteristics and utility of the version that addresses observed executive functioning in the earliest developmental stage (the BRIEF-P) and the extent to which the scale relates to disruptive behavior disorders are not yet known. For this reason, it would seem necessary to assess the scale further.

The questionnaire, which is quick and easy to answer, is made up of 63 items on five clinical scales: 1) Inhibit, reflecting problems in inhibitory control; 2) shift, reflecting difficulties in moving freely among situations, activities or aspects of a problem; 3) emotional control, covering difficulties for modulating emotional response; 4) working memory, measuring difficulties in holding information in mind for completing a task; and 5) plan/organize, assessing problems for anticipating future events and taking appropriate steps, and for putting information in order to achieve an objective. Principal components analysis of the five clinical scales on data for normative and clinical samples of the answers from parent or teachers yielded a three-factor model: 1) Inhibitory Self-Control Index (ISCI), made up of inhibit and emotional control; 2) Flexibility index (FI), made up of shift and emotional control, and 3)

Emergent Metacognition Index (EMI), made up of working memory and plan/organize, explaining 90% (parents) and 92% (teachers) of the variance (Gioia et al., 2003). The items of the five clinical scales also provide a Global Executive Composite (GEC). The questionnaire has two validity scales (negativity and inconsistency) that help consideration of the validity of the data provided. The simplicity of application of the BRIEF has made this questionnaire one of the most widely used in the assessment of executive functioning in clinical, educational and research settings.

In contrast to the growing body of recent literature on the 5-18 version (BRIEF), which has shown evidence that broadly supports its psychometric properties, few studies have focused on the characteristics of the BRIEF-P. Valid and reliable assessment tools that help to identify early in life the specific disabilities of these impairing disorders could be beneficial for the planning of interventions in children affected. Two studies have focused on the psychometric properties of the BRIEF-P. In the teacher's normative samples ( $N = 302$  2 to 5-year-old children and subgroups) the questionnaire has shown good reliability (Cronbach's alpha between .86 and .95), low parent-teacher agreement (overall mean correlation .19), good mean test-retest correlation in 2 to 5 weeks (mean correlation .83), and significant association with attention and externalizing behavior scales (Gioia et al., 2003). Working memory and plan/organize (and EMI) obtained the strongest correlations with inattention, and inhibit with hyperactivity. Bonillo, Araujo, Jané, Capdevila, and Riera (2012) carried out a second-order factor analysis of the 5 scales of the parent and teacher Catalan version for 3 to 6-year-old children and obtained a satisfactory 3-factor solution (emergent metacognition, flexibility, and inhibitory self-control) with very good reliability indexes, but they failed to confirm the complete structure (63 items in five scales and

the three broader indexes) in the two forms of the questionnaire (teachers and parents) through Confirmatory Factor Analysis because of the lack of fit.

The executive functions are a core component of ADHD (Barkley, 2012; Lambek et al., 2011). There is scarce literature regarding the usefulness or incremental validity of the BRIEF-P for identifying disruptive behavior disorders. Mahone and Hoffman (2007) compared twenty-five 3 to 5.9-year-old children with and without ADHD using the parents' BRIEF-P and found that all the scales and global indices significantly discriminated between groups, with the greatest effect observed for the Working memory scale. They also reported significant correlations between BRIEF-P indexes and Conners' Parent Rating Scale-Revised (Conners, 1997), which assesses attention deficit/hyperactivity disorder (ADHD) behaviors. In this line, Sherman and Brooks (2010) have noted some degree of overlap between the behaviors described in the BRIEF-P scales and those included in ADHD screening scales. Gioia et al. (2003) pointed out that although some specificity in the relationships between BRIEF-P scales and ADHD were found, in comparison to the school-age samples, the associations were less differentiated in the preschoolers. Therefore, more information is needed about the validity of the questionnaire in preschoolers.

The objective of this work is to evaluate the psychometric properties of the teacher BRIEF-P in preschool children at age 3, in an extensive Spanish community sample, and to report data about its usefulness for clinical assessment of disruptive behavior disorders. The specific goals are: (a) to confirm the five scales plus three global indices structures proposed by Gioia et al. (2003); (b) to analyze the internal consistency reliability of the derived factors; (c) to provide validity evidence in relation to external variables relevant in mental health, such as psychopathology, temperament, functional impairment, or executive functioning; (d) to analyze whether the BRIEF-P

adds significant information for the identification of disruptive behavior disorders when used in combination with other measures of behavior problems or performance-based executive functioning; and (e) to assess the screening accuracy of the BRIEF-P for identifying DSM-IV disruptive behavior disorders.

## Method

### *Participants*

Data are from a large-scale longitudinal study of behavioral problems in preschool children from age 3 (Ezpeleta, Osa, & Doménech, Submitted). The measures were applied in a two-phase design, starting with a random sample of 2,283 children selected from the census of preschoolers in grade P3 (3 years old) in Barcelona for 2009-10. A total of 1,341 families agreed to participate in the first phase of the study (58.7%). Of these, 33.6% were of high socioeconomic status, 43.1% middle and 23.3% low. Children's mean age was 3.0 years ( $SD = 0.18$ ), 683 were boys (50.9%), and 89.3% were white. There were no sex differences ( $p = .95$ ) between those who agreed to participate and those who declined, but semi-public schools were significantly more likely to refuse to participate than public ones ( $p < .001$ ), and high socioeconomic status (SES) families participated more than low status families ( $p < .001$ ). The parents of children participating in this first phase completed the Strengths and Difficulties Questionnaire (SDQ<sup>3-4</sup>) parents' version (Goodman, 1997), which was used for screening purposes.

In the second phase, all children with a positive screening score for behavioral problems and a random sample of 30% of children with a negative screening score were invited to continue. The final second-phase sample included families of 622 children (10.6% of those invited refused to participate in the second phase). No statistical



differences were found on comparing participants and refusals by sex ( $p = .82$ ) or by type of school ( $p = .85$ ). Children's mean age was 3.0 ( $SD = 0.16$ ), and 311 were boys (50%) and 88.9% were white, while 32.9% were of high socioeconomic status, 45.7% middle, and 21.4% low. Weighted DSM-IV prevalence in the final sample ( $N = 622$ ) was as follows: 3.6% of the children presented attention deficit/hyperactivity disorder (3.1% hyperactive and 0.5% combined), 6.9% ODD, 1.4% CD, 0.4% major depression, 2.2% separation anxiety, 3.7% specific phobia and 1.9% social phobia. A total of 17.6% of the mothers and 22.6% of the fathers had less than 7 years of education. BRIEF-P was unavailable for 2 children, resulting in a final sample of 620 for the analyses. At age four, 604 children remained in the study.

Children showing intellectual disability or pervasive developmental disorders, and families with difficulties with Spanish or Catalan, without a primary caregiver that could report about the child, or that were moving to another city within a year were excluded.

Ninety-four teachers from 54 schools answered the BRIEF-P (96.8 females, representative of the percentage of women working in infant education in our country). Participating teachers had known the 3-year-olds for a mean of 7.6 months ( $SD = 2.2$ ). In total, 3.4% of the teachers indicated that they had met the children less than 3 months previously, and 18.6% less than 6 months before. A comparison was made between BRIEF-P mean scores of all the scales (including the two validity scales) of teachers who had known the children for at least 6 months and those who had not, yielding no significant differences ( $p$  from .10 to .88). Teachers spent between 5 and 8 hours a day with the children at the school. Since there were no differences according to whether the children had been known to the teacher for more or less than 6 months, and the number of hours the teachers shared with the children guaranteed the possibility of good

observation, we retained in the analyses the assessments of all the teachers in the sample, so as to give more power to the statistical analyses.

### *Instruments*

The *Behavior Rating Inventory of Executive Function for Preschool Children* (BRIEF-P; Gioia et al., 2000), described previously, contains 63 items referring to behaviors in the last 6 months, and with 3 response options (0=“*never*” to 2=“*very often/always*”) reflecting the degree to which these behaviors are a problem. It takes approximately 10 to 15 minutes to complete. It was answered by teachers. The Spanish version provided by the publisher was used for the study.

Large differences in T-scores between the three BRIEF-P second-order-scales ISCI, FI and EMI (which preclude the use of GEC index as a summary measure) were found in 16.3% of the boys and 19.2% of the girls (differences by sex  $p = .350$ ). There were no differences in the percentage of inconsistent questionnaires for boys (3.2%) and girls (1.3%) ( $p = .114$ ), or in the percentage of high scores on the negative scale (2.2% for boys and 3.6% for girls,  $p = .322$ ).

The *Strengths and Difficulties Questionnaire* (SDQ<sup>3-4</sup>; Goodman, 1997) for parents of 3 to 4-year-old children was used as a screening instrument. The SDQ<sup>3-4</sup> is made up of 25 items with 3 response options (0: *not true*; 1: *somewhat true*; 2: *certainly true*) on five scales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior. Two broader internalizing (sum of the items of Emotional and Peers subscales) and externalizing factor (sum of Conduct and Hyperactivity subscales; A. Goodman, Lamping, & Ploubidis, 2010) were also used. The Cronbach's alpha reliability of the scales in the present sample ranged from

.55 (peers) to .76 (externalizing) for parents and between .63 (peers) and .84 (hyperactivity) for teachers (Ezpeleta, Granero, de la Osa, Penelo, & Domènech, 2012).

The *Diagnostic Interview for Children and Adolescents for Parents of Preschool and Young Children* (DICA-PPYC; Ezpeleta, Osa, Granero, Doménech, & Reich, 2011) is a semi-structured diagnostic interview designed to assess the most common psychological disorders at ages 3-7 years according to the DSM-IV-TR criteria (American Psychiatric Association, 2000). Parents were interviewed about the presence of symptoms, and where information on symptoms was available additional questions explored service use. Disorders were assessed over the lifetime.

The *Children's Behavior Questionnaire for ages 3-7 years* (CBQ3-7; (Rothbart, Ahadi, Hershey, & Fisher, 2001) measures reactive and self-regulative temperament, with 94 items on a 7-point Likert-type scale, ranging from 1 (*extremely untrue*) to 7 (*extremely true*). It was answered by parents. The three broad dimensions of temperament: negative affectivity (anger-frustration, discomfort, fear, sadness, soothability), effortful control (Attention Focusing, Inhibitory Control, Low-intensity Pleasure, Perceptual Sensitivity), and surgency (Activity Level, High-intensity Pleasure, Impulsivity, Shyness) were used for this study. Cronbach's alpha of the scales in the present sample ranged from .71 (negative affectivity) to .79 (effortful control).

The *Children's Global Assessment Scale* (CGAS; Ezpeleta, Granero, & de la Osa, 1999; Shaffer et al., 1983) was applied by the interviewer after the diagnostic interview. It was used to rate global functional impairment associated with psychopathology (0 indicating the highest impairment and 100 the lowest). Scores above 70 indicate normal functioning.

The *Kiddie-Continuous Performance Test* (K-CPT; Conners, 2006) is a performance-based measure of attention function and response inhibition for children

aged 4-5. The software present stimuli consisting in familiar pictures for very young children (doll, truck, ball, etc.), and response/no response are to be made depending on the picture that appears. The program provides 12 variables related to attention: *omission* means lack of a required response; *commission* is a response after a stimulus that requires none; *hit reaction time* is the mean response time over all 5 time blocks; *hit reaction time standard error* is a measure of erratic responding; *variability of standard errors* is the response time consistency; *attentiveness ( $d'$ )* denotes how well the individual discriminates between targets and nontargets; *perseverations* is a response that occurs less than 100ms following a stimulus; and the remaining variables are measures of reaction time. K-CPT was administered in the second year of the longitudinal study, when the children were 4 years old (as stipulated in the questionnaire manual).

### *Procedure*

The longitudinal project was approved by the ethics review committee of the authors' institution. The heads of the schools participating, as well as the children's parents, received a complete description of the study. Families were recruited at the schools and gave written consent. All parents of children from P3 (3-year-olds) in the participating schools were invited to answer the SDQ<sup>3-4</sup>. Families who agreed and met the screening criteria were contacted by telephone and interviewed at the school. Interviewers were previously trained and were blind to the children's screening group. After the interview, the interviewer completed the CGAS and parents answered the questionnaires. After obtaining permission from the parents, the teachers' questionnaires were given to them for completion before the end of the academic year.

All the measures were completed when the children were 3 years old, except the K-CPT, which was completed at age 4. Assessments were carried out with a one-year interval (mean between assessments 11.4 months, SD: 2.3).

### *Statistical analysis*

The statistical analysis was carried out with SPSS19 and Mplus6. Because of the multistage sample, data were analyzed through Complex Samples tools in SPSS, creating a plan file with sampling weights inversely proportional to the probability of participant selection, and with the case weighting procedure in the Mplus6 program.

The 63 BRIEF-P items were submitted to Confirmatory Factor Analysis (CFA) with Mplus6, using Weighted Least Squares Means and Variance (WLSMV) adjusted for the categorical data method of estimation. Covariance matrices were analyzed. The model considered was a 5-first-order-factor plus a 3-second-order-factor model, following the scoring instructions (Gioia et al., 2003). The 5-first-order factors considered were: Inhibit (IN, 16 items), Emotional Control (EC, 10 items), Shift (SH, 10 items), Working Memory (WM, 17 items), and Plan/Organize (PO, 10 items). The 3-second-order-factor model added: Inhibitory Self-Control Index (ISCI, 26 items grouping IN and EC), Flexibility Index (FI, 20 items grouping SH and also EC), and Emergent Metacognition Index (EMI, 27 items comprising WM and PO). Goodness of fit was assessed with the common fit indices (Jackson, Gillaspay, & Purc-Stephenson, 2009):  $\chi^2$ , Tucker and Lewis index (TLI), comparative fit index (CFI), and Root Mean Square Error of Approximation (RMSEA). The following thresholds were adopted:  $RMSEA < .05$  and  $TLI$  and  $CFI > .95$  are indicative of good fit, whereas  $RMSEA < .08$  and  $TLI$  and  $CFI > .90$  represent reasonable fit (Marsh, Hau, & Wen, 2004). Internal consistency of the derived scores was measured with Cronbach's  $\alpha$ .

The association between BRIEF-P and other psychological measures was carried out with the Pearson correlation;  $|R| \geq .30$  was considered a good effect size.

Hierarchical binary logistic regressions explored the incremental validity of BRIEF-P scale scores (added to the SDQ, CGAS and K-CPT scores) to discriminate the presence of attention-deficit/hyperactivity disorder (ADHD), oppositional defiant disorder (ODD) and disruptive disorders (ADHD, ODD or conduct disorder – CD). The incremental validity was measured with the change/increase of Nagelkerke's  $R^2$  coefficient ( $\Delta R^2$ ), comparing the first block including the SDQ, CGAS or K-CPT and the second block which added the BRIEF-P scores.

Binary logistic regressions valued the screening accuracy of BRIEF-P to identify ADHD and ODD at ages 3 and 4. The area under the ROC curve (AUC) measured the global discriminative accuracy of the BRIEF-P scores.

The best empirical cut-off of each BRIEF-P scale was calculated to detect the presence of ADHD, ODD, and disruptive behavior disorders. Given that we were analyzing here the supposed use of the questionnaire for screening purposes, the selection was based on sensitivity of no less than .70.

## Results

### *Confirmatory Factor Analysis for BRIEF-P items and Internal Consistency*

Table 1 shows the results for the model considered. To adjust the model correctly, four items showing an extreme endorsement were removed: item 1 (“overreacts to small problems”), 13 (“has to be more closely supervised than similar playmates”), 35 (“has trouble changing activities”) and 58 (“has trouble remembering something, even after a brief period of time”). The 5-first-order-factor-model plus a 3-second-order-factor model yielded acceptable goodness-of-fit indices:  $\chi^2(1646) =$

4892.6; CFI = .89; TLI = .89; RMSEA = .056. All factor loadings and factor correlations were statistically significant, and standardized factor loadings were above .50. Internal consistency was satisfactory ( $\alpha \geq .87$ ) for the resulting scales.

The following analyses are conducted with the direct scores derived, and therefore items 1, 13, 35 and 58 are not included.

#### **INSERT TABLE 1**

##### *Convergent Validity of BRIEF-P with other psychological measures*

Table 2 shows the correlations of the BRIEF-P with other psychological measures. Only BRIEF-P-Inhibit correlated positively with SDQ-parents externalizing problems. However, most of the correlation coefficients between BRIEF-P and SDQ-teachers scales showed significant positive associations. CBQ-surgency correlated positively with BRIEF-P-inhibit and negatively with BRIEF-P-shift, while CBQ-effortful control correlated negatively with BRIEF-P-working-memory, plan-organize, EMI, and GEC. None of the BRIEF-P scales was associated with the performance-based measure of attention (K-CPT).

#### **INSERT TABLE 2**

##### *Does the BRIEF-P add significant information for identifying disruptive behavior disorders when used in combination with other measures of behavior problems or performance-based executive functioning?*

The results of the hierarchical binary logistic regressions that valued the incremental validity of BRIEF-P to identify the presence of disruptive disorders (Table 3) indicate that BRIEF-P achieved incremental predictive accuracy to detect ADHD added to SDQ-parents ( $\Delta R^2 = 5.3\%$ ), CGAS ( $\Delta R^2 = 11.4\%$ ) and K-CPT ( $\Delta R^2 = 15.3\%$ ). The specific BRIEF-P scale scores that obtained significant incremental predictive

accuracy for detecting the presence of ADHD were Inhibit ( $\Delta R^2 = 4.9\%$  added to SDQ-parents and  $\Delta R^2 = 11.0\%$  added to CGAS) and Inhibit plus Shift ( $\Delta R^2 = 13.3\%$ ). No significant incremental validity was found when the BRIEF-P measures were added to SDQ-teachers scores for detecting the presence of ADHD. The BRIEF-P scales only achieved significant incremental predictive accuracy for the presence of ODD added to K-CPT, but the increase was low ( $\Delta R^2 = 1.9\%$  for the whole BRIEF-P questionnaire and  $\Delta R^2 = 0.9\%$  for Emotional Control). Considering the presence of any disruptive disorder (ADHD, ODD or CD), the BRIEF-P achieved significant incremental validity added to K-CPT ( $\Delta R^2 = 5.6\%$  for the complete questionnaire and  $\Delta R^2 = 3.2\%$  for the specific scale with significant predictive association, Emotional Control) (Table 3 online).

### INSERT TABLE 3

New hierarchical regressions in three steps were carried out, so as to value the incremental validity of BRIEF-P (entered in the third step) on SDQ-parents (entered in the first-step) and K-CPT (entered on the second step). BRIEF-P achieved significant predictive increase for the presence of ADHD ( $\Delta R^2 = 6.6\%$  for the complete BRIEF-P questionnaire and  $\Delta R^2 = 5.5\%$  for Inhibit) and ODD ( $\Delta R^2 = 2.0\%$  for the complete BRIEF-P questionnaire and  $\Delta R^2 = 0.8\%$  for Inhibit) (Table 3 online).

#### *Screening accuracy of BRIEF-P for identifying DSM-IV disorders*

Table 4 (top) shows the means of BRIEF-P scale scores for children with and without the presence of ADHD diagnosis and the binary logistic regressions that valued the screening accuracy of the questionnaire to identify this disorder. The BRIEF-P achieved acceptable discriminative accuracy for ADHD: AUC ranged from .74 to .80, and Nagelkerke's  $R^2$  was between 11.0% and 18.5%. The specific BRIEF-P scales with significant predictive accuracy were inhibit, ISCI, GEC at ages 3 and 4, and FI at age 4.



The discriminative accuracy for identifying the presence of ODD was low (AUC between .55 and .62 and  $R^2$  between 0.2% and 5.3%) and non-significant for the second order scales ISCI, FI and EMI (Table 4, bottom). The only scales with significant discriminative accuracy for the screening of ODD were emotional control and GEC.

#### **INSERT TABLE 4**

Table 5 shows the empirical best cut-off of BRIEF-P scale scores for the screening of ADHD, ODD and disruptive behavior disorders (ADHD, ODD or CD).

#### **INSERT TABLE 5**

#### *Description of scores by sex*

Annex 1 (online material) shows the means and standard deviations of raw scores for the BRIEF-P scales. Some statistical differences emerged by sex ( $p < .05$ ; boys scored higher than girls in inhibit, working memory, plan-organize and ISCI, EMI and GEC indexes). Effect sizes comparing mean scores by sex through Cohen's  $d$  (Cohen, 1988) were  $< .50$ , indicating that the effects were small.

Annex 2 (online material) shows the clinical thresholds for direct scores (percentiles 80, 90, and 95).

#### **Discussion**

This is one of the few studies analyzing the psychometric properties of the teachers' BRIEF-P. The 5-first-order-factor plus 3-second-order-factor model was confirmed with satisfactory internal consistency, moderate convergent validity with other measures of psychopathology and temperament, and good ability to discriminate between those children with ADHD and those without. BRIEF-P scales were not associated with a performance-based measure of attention and inhibition. The teacher's

BRIEF-P added a considerable percentage of significant clinical information when used with other instruments that assess psychopathology reported by parents, functional impairment or performance-based attention and inhibition, showing that it might be a useful assessment tool, specifically for ADHD.

Regarding validity, it was possible to confirm the internal structure proposed by Gioia et al. (2003). Different alternative models were considered, a model containing the 63 items and 5 first-order factors, and another containing the 63 items and 3 first-order factors, but they yielded unsatisfactory solutions. Therefore, the 5-first-order-factor plus a 3-second-order-factor model emerged as most appropriate for explaining the structure underlying the items of the BRIEF-P, despite the fact that four items were omitted for fitting the model (two in inhibit, one in emotional control and one in shift). This deletion could limit slightly the comparisons with results of other studies using this questionnaire. However, the results of the psychometric properties reported throughout this work with the direct scores of the empirical factors support the reliability and validity of the structure proposed. To our knowledge this is the first study in preschoolers that obtained this internal structure by factorizing items instead of scales.

In relation to external measures, there was strong convergence between the BRIEF-P and measures of psychopathology (SDQ<sup>3-4</sup>) when the informant was the same (teacher), but the association was low when the information from different informants (parent-teacher) was correlated. Teachers were consistent in reporting the behavior described in the two different instruments, and reported that behaviors reflecting dysfunction in executive functioning were positively correlated with behaviors reflecting psychopathology. The lack of agreement between teachers' and parents' reports is a common phenomenon in psychological assessment, which has been explained by the fact that different informants observe different behaviors in different

contexts (de los Reyes & Kazdin, 2005; Smith, 2007). Our results highlight the need to include both teacher and parent ratings in the assessment of disruptive behavior disorders in preschool children. Also, there was a significant negative association between BRIEF-P working memory and plan-organize and the temperament dimension effortful control, typically associated with ADHD (Martel & Nigg, 2006), and between surgency (extraversion) and difficulties in BRIEF-P inhibit and low shift, supporting convergence with temperament measures.

Also in relation to convergent validity, a lack of association has been observed between traditional performance-based measures of executive functioning and questionnaires in adults and in children, to the extent that researchers have asked themselves whether the BRIEF-P is more a measure of impairment than a measure of executive function (Bakar, Taner, Soysal, Karakas, & Turgay, 2011; Garcia-Molina, Tormos, Bernabeu, Junque, & Roig-Rovira, 2012). We found no significant association between BRIEF-P scales and performance-based measures of attention and inhibition, in line with the previous literature on preschoolers and older children (Mahone & Hoffman, 2007; Toplak, Bucciarelli, Jain, & Tannock, 2009). However, most interestingly, we extended our study to incremental validity, and also studied the amount of variance explained by the BRIEF-P when used in combination with other measures. The incremental validity of the BRIEF-P has not been tested previously in preschoolers. We found that the BRIEF-P added a substantial amount of information for identifying ADHD to questionnaires of psychopathology answered by parents, to diagnostic interview with parents for rating functional impairment, and to K-CPT performance-based measures of attention and inhibition answered by the child. Used with these measures and these reporters, the BRIEF-P answered by teachers contributes specific and significant clinical information on dysfunction in executive functioning.

However, when the teachers answered a measure of psychopathology (such as SDQ<sup>3-4</sup> conduct and hyperactivity scales) and the BRIEF-P, the incremental validity of the BRIEF-P was negligible, and in this case the results yielded by examination of everyday behavior for assessing executive function as observed in school are very close to those yielded by the observation of psychopathological behavior. Consequently, different informants must be used in order to obtain the incremental information with the BRIEF-P for identifying ADHD when used together with other measures of psychopathology.

In the area of test-criterion relationships, the BRIEF-P also showed good discriminative ability to identify cases with and without ADHD cross-sectionally and one year later. The most predictive scale was Inhibit, and the best index was ISCI. Considering that in our sample the most frequent ADHD subtype was hyperactive, these results are congruent with executive dysfunction found in this disorder (Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). In contrast, the discriminative ability to identify ODD was poor. Several reports have indicated that deficits in executive functioning are more specific of ADHD than of ODD (Brocki, Nyberg, Thorell, & Bohlin, 2007; Thorell & Wahlstedt, 2006), and that ADHD is associated with inhibition problems, whereas ODD is more related to emotional dysregulation (Espy, Sheffield, Wiebe, Clark, & Moehr, 2011).

Finally, we also provide different cut-offs that might be useful for identifying disruptive behavior problems if the questionnaire is used for screening purposes. If executive dysfunction were to be identified early in life, preventive measures could be implemented. Early detection is especially relevant in the general population, as used in the present study.

Some limitations should be taken into account on interpreting the present results. Since parents' ratings on the BRIEF-P were not obtained, the findings cannot be

generalized to the use of the questionnaire by parents. Also, we studied only 3 and 4-year-old children, and this limits the generalization of the results to other ages.

However, we focused the study on this age group because it corresponds to the start of schooling, when the detection of dysfunctions for prevention is most important. The methods used in this study were standard in Western clinical psychology, as far as diagnostic criteria (DSM-IV-TR) and assessment techniques are concerned, so that the constructs assessed are defined similarly in other Western cultures. However, despite the fact that the internal structure and internal consistency we found are similar to those presented in other published studies with samples of other ages and from other countries, teachers may evaluate the behaviors differently across different cultures. Consequently, the present results may not be generalizable to other cultures without the support of accumulated empirical evidence guaranteeing the equivalence of the constructs and their correlates. We obtained a higher percentage of differences in T-Scores between the three BRIEF-P second-order-scale scores than that estimated in the manual (10%), so that caution should be exercised on interpreting the GEC as a summary executive index measure. Furthermore, in a small percentage of cases there were inconsistencies in the answers to the questionnaire, as measured by the inconsistency scale. We should bear in mind the questionnaire authors' indication that the content of the inconsistency scale items, though very similar, is not exactly the same. Analysis of the inconsistencies found indicates that 80% of them were between item 48 ("Talks or plays too loudly") and 54 ("Has trouble putting the brakes on his/her actions even after being asked"), and between 1 ("Overreacts to small problems") and 26 ("Small events trigger big reactions"), where the interpretation of the teachers might have been different. Another possible limitation is that because we studied a sample of the general population and psychopathology is not very frequent in community samples,

the discriminative power may have been underestimated. And finally, few families of low socioeconomic status participated, and this could have led to bias, insofar as our results might be better representing executive dysfunction in more favored socioeconomic classes.

## References

- American Psychiatric Association. (2000). *DSM-IV Diagnostic and statistical manual of mental disorders* (4th Text Revised ed.). Washington, DC: American Psychiatric Press.
- Bakar, E. E., Taner, Y. I., Soysal, A. S., Karakas, S., & Turgay, A. (2011). Behavioral Rating Inventory and Laboratory Tests Measure Different Aspects of Executive Functioning in Boys: A Validity Study. *Klinik Psikofarmakoloji Bulteni-Bulletin of Clinical Psychopharmacology*, *21*(4), 302-316.
- Barkley, R. A. (2012). *Executive Functions. What They Are, How They Work, and Why They Evolved* New York: Guilford.
- Bonillo, A., Araujo, E. A., Jané, M. C., Capdevila, C., & Riera, R. (2012). Brief Report: Validation of Catalan version of BRIEF-P. *Child Neuropsychology*, *18*, 344-355.
- Brocki, K. C., Nyberg, L., Thorell, L. B., & Bohlin, G. (2007). Early concurrent and longitudinal symptoms of ADHD and ODD: relations to different types of inhibitory control and working memory. *Journal of Child Psychology and Psychiatry*, *48*(10), 1033-1041.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale (NJ): Lawrence Earlbaum Associates.
- Conners, C. K. (1997). *Conners' rating scales revised*. Niagara Falls, NY: Multi-Health Systems, Inc.
- Conners, C. K. (2006). *Kiddie Conner's Continuous Performance Test technical guide and software manual*. Toronto: Multi-Health Systems.
- de los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, *131*, 483-509.

- Espy, K. A., Sheffield, T. D., Wiebe, S. A., Clark, C. A. C., & Moehr, M. J. (2011). Executive control and dimensions of problem behaviors in preschool children. *Journal of Child Psychology and Psychiatry*, 52(1), 33-46.
- Ezpeleta, L., Granero, R., & de la Osa, N. (1999). Evaluación del deterioro en niños y adolescentes a través de la Children's Global Assessment Scale (CGAS). *Revista de Psiquiatría Infanto-Juvenil*, 1, 18-26.
- Ezpeleta, L., Granero, R., Osa, N. d. l., Penelo, E., & Doménech, J. M. (in press). Psychometric properties of the Strengths and Difficulties Questionnaire3-4 in 3 year-old preschoolers. *Comprehensive Psychiatry*.
- Ezpeleta, L., Osa, N. d. l., & Doménech, J. M. (Submitted). Prevalence of DSM-IV disorders, comorbidity and impairment in 3-year-old Spanish preschoolers.
- Ezpeleta, L., Osa, N. d. l., Granero, R., Doménech, J. M., & Reich, W. (2011). The Diagnostic Interview for Children and Adolescents for Parents of Preschool Children. *Psychiatry Research*, 190, 137-144.
- Garcia-Molina, A., Tormos, J. M., Bernabeu, M., Junque, C., & Roig-Rovira, T. (2012). Do traditional executive measures tell us anything about daily-life functioning after traumatic brain injury in Spanish-speaking individuals? *Brain Injury*, 26(6), 864-874.
- Gioia, G. A., Espy, K. A., & Isquith, P. K. (2003). *Behavior Rating Inventory of Executive Function - Preschool version*. Lutz, FL: Psychological Assessment Resources.
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *BRIEF-P Behavior Rating Inventory of Executive Function- Preschool Version (Inventario de Evaluación Conductual de las Funciones Ejecutivas - Versión Preescolar)* (L. Ezpeleta, Trans.). Lutz, FL: Psychological Assessment Resources, Inc.



- Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the strengths and difficulties questionnaire (SDQ): Data from british parents, teachers and children. *Journal of Abnormal Child Psychology*, 38(8), 1179-1191.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 581-586
- Jackson, D. L., Gillaspay, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychological Methods*, 14, 6-23.
- Lambek, R., Tannock, R., Dalsgaard, S., Trillingsgaard, A., Damm, D., & Thomsen, P. H. (2011). Executive Dysfunction in School-Age Children With ADHD. *Journal of Attention Disorders*, 15(8), 646-655.
- Mahone, E. M., & Hoffman, J. (2007). Behavior ratings of executive function among preschoolers with ADHD. *Clinical Neuropsychologist*, 21(4), 569-586.
- 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 Equations Modeling*, 11, 320-341.
- Martel, M. M., & Nigg, J. T. (2006). Child ADHD and personality/temperament traits of reactive and effortful control, resiliency, and emotionality. *Journal of Child Psychology and Psychiatry*, 47(11), 1175-1183.
- Rothbart, M. K., Ahadi, S. A., Hershey, K. L., & Fisher, P. (2001). Investigations of temperament at 3-7 years: The Children's Behavior Questionnaire. *Child Development*, 72, 1394-1408.

- Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., et al. (1983). A Children's Global Assessment Scale (CGAS). *Archives of General Psychiatry*, *40*, 1228-1231.
- Sherman, E. M. S., & Brooks, B. L. (2010). Behavior Rating Inventory of Executive Function - Preschool Version (BRIEF-P): Test Review and Clinical Guidelines for Use. *Child Neuropsychology*, *16*(5), 503-519.
- Smith, S. R. (2007). Making sense of multiple informants in child and adolescent psychopathology - A guide for clinicians. *Journal of Psychoeducational Assessment*, *25*(2), 139-149.
- Thorell, L. B., & Wahlstedt, C. (2006). Executive functioning deficits in relation to symptoms of ADHD and/or ODD in preschool children. *Infant and Child Development*, *15*(5), 503-518.
- Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2009). Executive Functions: Performance-Based Measures and the Behavior Rating Inventory of Executive Function (BRIEF) in Adolescents with Attention Deficit/Hyperactivity Disorder (ADHD). *Child Neuropsychology*, *15*(1), 53-72.
- 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.

Table 1. *Confirmatory Factor Analysis for the BRIEF-P: Standardized Solution.*

	First-order factors					Second-order factors		
	IN	EC	SH	WM	PO	ISCI	FI	EMI
$\lambda$ (1-order factors)								
IN03 Unaware of how behavior affects others	.626							
IN08 Does not stop laughing	.796							
IN13 Closely supervised	---							
IN18 Acts wilder	.937							
IN23 Fidgety	.842							
IN28 Impulsive	.832							
IN33 Unaware when causes negative reactions	.862							
IN38 Unaware that certain actions bother others	.866							
IN43 Out of control	.939							
IN48 Too loud	.767							
IN52 Too wild	.942							
IN54 Trouble putting brakes on	.907							
IN56 Completes tasks too quickly	.538							
IN58 Easily sidetracked	---							
IN60 Becomes silly	.851							
IN62 Careless	.785							
EC01 Overreacts		---						
EC06 Explosive		.974						
EC11 Upset easily		.675						
EC16 Easy outbursts		.814						
EC21 Mood changes		.833						
EC26 Small events trigger big reactions		.801						
EC31 Intense angry		.760						
EC36 Reacts strongly		.869						
EC41 Overwhelmed		.859						
EC46 Disappointed for long time		.600						
SH05 Upset with new situations			.588					
SH10 Trouble adjusting to new people			.806					
SH15 Upset by change			.903					
SH20 Takes long time to adapt in new situations			.906					
SH25 Bothered by loud noises			.748					
SH30 Disturbed by changes			.798					
SH35 Trouble changing activities			---					
SH40 Trouble "joining in" in unfamiliar situations			.765					
SH45 Resists change			.911					
SH50 Overwhelmed in crowded places			.992					
WM02 Remembers only the first or last				.812				
WM07 Trouble complete tasks				.862				
WM12 Trouble concentrating				.889				
WM17 Repeats mistakes				.886				
WM22 Silly mistakes				.835				
WM27 Trouble with tasks of more than one step				.900				
WM32 Needs help to stay on task				.857				
WM37 Forgets in the middle of an activity				.868				
WM42 Unfinished tasks				.922				
WM47 Trouble staying on topic				.676				
WM51 Trouble starting activities				.866				
WM53 Does not try according to ability				.805				
WM55 Unable to finish a description				.721				
WM57 Unaware of own performance				.905				
WM59 Trouble remembering				.831				
WM61 Short attention span				.861				
WM63 Unaware of right or wrong performance				.910				
PO04 Disorganized					.785			
PO09 Needs reminding to begin tasks					.718			
PO14 Forgets orders					.829			
PO19 Cannot find things					.838			
PO24 Trouble following routines					.816			
PO29 Trouble thinking differently					.739			
PO34 Leaves messes					.821			
PO39 Caught up in small details					.782			
PO44 Cannot find things					.897			
PO49 Does not complete tasks					.908			
$\lambda$ (2-order factors)								
Inhibition						.857		
Emotional control						.511	.516	
Shift							.856	
Working-memory								.964
Plan-organize								.997
$\alpha$ (1-order factors)								
Inhibition	.93 (14)							
Emotional control		.88 (9)						
Shift			.87 (9)					
Working-memory				.95 (17)				
Plan-organize					.89 (10)			
$\phi$ and $\alpha$ (2-order factors)								
ISCI index						.94 (23)		
FI index						.376	.91 (18)	
EMI index						.710	.438	.96 (27)
Total GEC								.96 (59)

( $N = 620$ ). Standardized factor loadings ( $\lambda$ ), factor correlations ( $\phi$ ), and internal consistency ( $\alpha$ ).

IN: Inhibit; EC: Emotional Control; SH: Shift; WM: Working Memory; PO: Plan/Organize; ISCI: Inhibitory Self-Control Index (IN + EC); FI: Flexibility Index (SH + EC); EMI: Emergent Metacognition Index (WM + PO); GEC: Global Executive Composite

In italics: Cronbach's  $\alpha$  (no. items).

Table 2.

*Association Between BRIEF-P and Other Psychological Measures.*

	BRIEF-P scales →	IN	SH	EC	WM	PO	ISCI	FI	EMI	GEC
SDQ-Parents	Emotional symptoms	-.06	.19	.01	.09	.08	-.04	.11	.09	.07
	Conduct problems	.22	.02	.21	.20	.20	.25	.14	.20	.23
	ADHD	.29	-.07	.08	.25	.22	.24	.00	.25	.24
	Peers	.06	.14	.06	.18	.14	.07	.12	.17	.15
	Pro-social	.07	.06	.06	.09	.10	.08	.06	.09	.10
	Internalizing problems	.00	.20	.04	.17	.14	.02	.14	.16	.13
	Externalizing problems	<b>.32</b>	-.04	.16	.28	.26	.29	.07	.28	.28
	Total	.24	.07	.14	.29	.26	.23	.13	.29	.28
SDQ-Teachers	Emotional symptoms	.17	<b>.63</b>	<b>.49</b>	<b>.35</b>	<b>.35</b>	<b>.32</b>	<b>.64</b>	<b>.36</b>	<b>.46</b>
	Conduct problems	<b>.65</b>	.19	<b>.65</b>	<b>.30</b>	<b>.31</b>	<b>.73</b>	<b>.49</b>	<b>.31</b>	<b>.55</b>
	ADHD	<b>.71</b>	.12	<b>.35</b>	<b>.69</b>	<b>.64</b>	<b>.65</b>	.27	<b>.69</b>	<b>.71</b>
	Peers	.18	<b>.40</b>	.28	<b>.39</b>	<b>.37</b>	.24	<b>.39</b>	<b>.40</b>	<b>.40</b>
	Pro-social	<b>.41</b>	.21	<b>.33</b>	<b>.37</b>	<b>.38</b>	<b>.42</b>	<b>.31</b>	<b>.38</b>	<b>.45</b>
	Internalizing problems	.21	<b>.62</b>	<b>.47</b>	<b>.44</b>	<b>.43</b>	<b>.34</b>	<b>.62</b>	<b>.45</b>	<b>.52</b>
	Externalizing problems	<b>.81</b>	.17	<b>.56</b>	<b>.62</b>	<b>.60</b>	<b>.81</b>	<b>.43</b>	<b>.63</b>	<b>.76</b>
	Total	<b>.69</b>	<b>.44</b>	<b>.64</b>	<b>.68</b>	<b>.65</b>	<b>.76</b>	<b>.62</b>	<b>.68</b>	<b>.82</b>
CBQ	Surgency	<b>.30</b>	<b>-.32</b>	.06	.04	.04	.24	-.14	.05	.08
	Negative Affectivity	-.11	.17	.07	-.02	.00	-.05	.13	-.01	.00
	Effortful control	-.27	-.03	-.11	<b>-.36</b>	<b>-.34</b>	-.24	-.08	<b>-.36</b>	<b>-.32</b>

Pearson's correlation coefficients. In bold:  $|R| \geq .30$

IN: Inhibit; SH: Shift; EC: Emotional Control; WM: Working Memory; PO: Plan/Organize; ISCI: Inhibitory Self-Control Index (IN + EC); FI: Flexibility Index (SH + EC); EMI: Emergent Metacognition Index (WM + PO); GEC: Global Executive Composite

Table 3.

*Incremental Validity of BRIEF-P to Identify ADHD and ODD.*

SDQ-parents	Criteria (dependent variable) →	ADHD		ODD		
		p	OR	p	OR	
First step	SDQ-parents: Conduct problems	<.001	1.47	<.001	1.62	
	SDQ-parents: ADHD	<.001	1.60	.216	1.10	
		$R^2 = .299$		$R^2 = .167$		
Second step	SDQ-parents: Conduct problems	<.001	1.44	<.001	1.64	
	SDQ-parents: ADHD	<.001	1.52	.082	1.14	
	BRIEF: Inhibit	<b>.037</b>	<b>1.08*</b>	.071	0.95	
	BRIEF: Shift	.641	0.96	.841	0.98	
	BRIEF: Emotional control	.333	1.06	.101	1.09	
	BRIEF: Working memory	.558	1.03	.410	0.96	
	BRIEF: Plan-organize	.658	0.95	.705	1.04	
			$\Delta R^2 = .053$		$\Delta R^2 = .018$	
		$Total R^2 = .352$		$Total R^2 = .185$		
SDQ-teachers		p	OR	p	OR	
First step	SDQ- teachers: Conduct problems	.011	1.25	.009	1.22	
	SDQ- teachers: ADHD	<.001	1.32	.597	0.97	
		$R^2 = .169$		$R^2 = .028$		
Second step	SDQ- teachers: Conduct problems	.196	1.16	.016	1.31	
	SDQ- teachers: ADHD	.085	1.19	.503	1.05	
	BRIEF: Inhibit	.397	1.05	.059	0.92	
	BRIEF: Shift	.219	0.90	.484	0.95	
	BRIEF: Emotional control	.525	1.05	.563	1.03	
	BRIEF: Working memory	.469	1.04	.872	0.99	
	BRIEF: Plan-organize	.867	0.98	.596	1.04	
			$\Delta R^2 = .016$		$\Delta R^2 = .017$	
		$Total R^2 = .185$		$Total R^2 = .045$		
CGAS		p	OR	p	OR	
First step	CGAS: total score	<.001	0.82	<.001	0.82	
		$R^2 = .351$		$R^2 = .361$		
Second step	CGAS: total score	<.001	0.82	<.001	0.82	
	BRIEF: Inhibit	<b>&lt;.001</b>	<b>1.15</b>	.377	0.97	
	BRIEF: Shift	.357	0.92	.283	0.93	
	BRIEF: Emotional control	.843	1.01	.069	1.11	
	BRIEF: Working memory	.509	1.04	.320	0.94	
	BRIEF: Plan-organize	.861	0.98	.513	1.07	
			$\Delta R^2 = .114$		$\Delta R^2 = .016$	
		$Total R^2 = .465$		$Total R^2 = .377$		
K-CPT						
First step	Omissions	.163	0.95	.684	1.01	
	Commissions	.179	0.94	.472	1.02	
	Hit Reaction Time - Overall	.108	0.94	.362	.98	
	Hit Reaction Time Standard Error - Overall	.001	1.22	.815	1.01	
	Variability of Standard Error	.447	0.97	.759	1.01	
	Detectability (d') (Attention)	.899	1.01	.172	.96	
	Response Style (β)	.667	1.01	.600	.99	
	Perseverations %	.804	1.00	.957	1.00	
	Hit RT Block Change	.397	1.01	.180	1.02	
	Hit SE Block Change	.830	0.99	.193	.98	
	Hit RT ISI Change	.925	1.00	.428	.98	
	Hit SE ISI Change	.366	1.02	.813	1.00	
			$R^2 = .088$		$R^2 = .020$	
	Second step	Omissions	.084	0.94	.726	1.01
Commissions		.189	0.93	.407	1.03	
Hit Reaction Time - Overall		.508	0.97	.390	0.98	
Hit Reaction Time Standard Error - Overall		.041	1.15	.835	1.01	
Variability of Standard Error		.477	0.96	.833	1.01	
Detectability (d') (Attention)		.924	1.00	.134	0.96	
Response Style (β)		.909	1.00	.573	0.99	
Perseverations %		.780	1.01	.945	1.00	
Hit RT Block Change		.358	1.02	.173	1.02	
Hit SE Block Change		.889	1.00	.221	0.98	
Hit RT ISI Change		.944	1.00	.473	0.98	
Hit SE ISI Change		.076	1.06	.800	1.00	
BRIEF: Inhibit		<b>.001</b>	<b>1.12</b>	.223	0.96	
BRIEF: Shift		<b>.044</b>	<b>0.86</b>	.336	0.93	
BRIEF: Emotional control		.089	1.09	<b>.017</b>	<b>1.13</b>	
BRIEF: Working memory	.133	1.08	.865	0.99		
BRIEF: Plan-organize	.696	0.96	.627	1.04		
		$\Delta R^2 = .153$		$\Delta R^2 = .019$		
		$Total R^2 = .241$		$Total R^2 = .039$		

R<sup>2</sup>: Nagelkerke's R<sup>2</sup> coefficient.  $\Delta R^2$ : Increase/change in the Nagelkerke's R<sup>2</sup> coefficient for the second step. In bold significant parameter

Table 4.

*Screening Accuracy of BRIEF-P to Identify ADHD and ODD.*

	ADHD at 3 years old						ADHD at 4 years old					
	Means for BRIEF		Binary logistic				Means for BRIEF		Binary logistic			
	ADHD Absent (n = 589)	ADHD Present (n = 33)	p	OR	IC 95% (OR)		ADHD Absent (n = 569)	ADHD Present (n = 35)	p	OR	IC 95% (OR)	
<b>BRIEF: first-order scales</b>												
Inhibit	19.8	27.0	<b>.004</b>	<b>1.11*</b>	<b>1.03</b>	<b>1.19</b>	19.7	26.7	<b>.001</b>	<b>1.11*</b>	<b>1.04</b>	<b>1.19</b>
Shift	12.0	12.2	.133	0.88	0.74	1.04	12.0	12.1	.135	0.88	0.74	1.04
Emotional control	11.9	14.6	.093	1.09	0.99	1.21	11.9	13.9	.481	1.04	0.94	1.15
Working memory	22.4	28.5	.209	1.06	0.97	1.16	22.2	28.9	.213	1.08	0.96	1.21
Plan-organize	13.3	16.4	.813	0.98	0.78	1.19	13.2	16.6	.908	0.99	0.82	1.20
				<b>AUC=.799*</b>						<b>R<sup>2</sup>=.167*</b>		
							<b>AUC=.792*</b>			<b>R<sup>2</sup>=.185*</b>		
<b>BRIEF: second-order scales</b>			p	OR	IC 95%				p	OR	IC 95%	
ISCI index	31.7	41.6	<b>&lt;.001</b>	<b>1.13*</b>	<b>1.07</b>	<b>1.19</b>	31.6	40.6	<b>&lt;.001</b>	<b>1.12*</b>	<b>1.06</b>	<b>1.19</b>
FI index	23.8	26.9	.030	0.92	0.86	.99	23.9	26.0	<b>.042</b>	<b>0.90*</b>	<b>0.82</b>	<b>1.00</b>
EMI index	35.7	44.9	.324	1.02	0.98	1.07	35.5	45.5	.038	1.04	1.00	1.08
				<b>AUC=.800*</b>						<b>R<sup>2</sup>=.160</b>		
							<b>AUC=.792*</b>			<b>R<sup>2</sup>=.179*</b>		
<b>BRIEF: total</b>			p	OR	IC 95%				p	OR	IC 95%	
GEC index	79.8	98.9	<b>&lt;.001</b>	<b>1.04*</b>	<b>1.03</b>	<b>1.06</b>	79.0	98.3	<b>&lt;.001</b>	<b>1.04*</b>	<b>1.03</b>	<b>1.06</b>
				<b>AUC=.760*</b>						<b>R<sup>2</sup>=.110*</b>		
							<b>AUC=.741*</b>			<b>R<sup>2</sup>=.121*</b>		

Se: sensitivity; Sp: specificity; AUC: Area Under ROC curve; R<sup>2</sup>: Nagelkerke's R<sup>2</sup>. Bold: significant coefficient (.05 level).

Table 5.

*Empirical Cut-off of BRIEF-P to Detect ADHD, ODD, and disruptive behavior disorders*

BRIEF-P scale	ADHD			ODD			Disruptive behavior disorders		
	Cut-off	Se	Sp	Cut-off	Se	Sp	Cut-off	Se	Sp
Inhibit	17	72.9	40.1	23	72.7	71.5	18	71.3	47.1
Shift	10	70.0	30.3	10	77.3	30.4	10	72.4	29.1
Emotional Control	11	71.7	45.5	11	81.8	44.7	11	73.6	43.2
Working Memory	19	73.3	39.5	21	72.7	57.3	20	71.3	48.0
Plan-Organize	12	71.7	40.5	13	72.7	67.8	12	75.9	42.1
ISCI index	28	71.7	40.5	34	72.7	67.8	30	71.3	49.9
FI index	21	75.0	35.1	21	78.3	34.5	21	74.7	33.3
EMI index	30	73.3	37.3	34	77.3	58.0	31	71.3	42.0
GEC index	73	71.7	47.6	82	77.3	65.5	75	71.3	49.9

Se: sensitivity; Sp: specificity; AUC: Area Under ROC curve.

Disruptive behavior disorders include attention deficit-hyperactivity disorder, oppositional-defiant disorder and conduct disorder