

Psychometric properties and clinical utility of the executive function inventory for children and adolescents: a large multistage populational study including children with ADHD

Marco Antônio Arruda, Renato Arruda & Luis Anunciação

To cite this article: Marco Antônio Arruda, Renato Arruda & Luis Anunciação (2020): Psychometric properties and clinical utility of the executive function inventory for children and adolescents: a large multistage populational study including children with ADHD, *Applied Neuropsychology: Child*, DOI: [10.1080/21622965.2020.1726353](https://doi.org/10.1080/21622965.2020.1726353)

To link to this article: <https://doi.org/10.1080/21622965.2020.1726353>



Published online: 02 Mar 2020.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



Psychometric properties and clinical utility of the executive function inventory for children and adolescents: a large multistage populational study including children with ADHD

Marco Antônio Arruda^a, Renato Arruda^b, and Luis Anunciação^c 

^aGlia Institute, Ribeirão Preto, Brazil; ^bDepartment of Neuroscience and Behavioral Sciences, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil; ^cDepartment of Psychometrics, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

ABSTRACT

Executive functions (EF) are a set of high order mental abilities that regulate cognition, emotions, and behavior. This study aims to report the construction and validation of a rating scale instrument for EF in children and adolescents aged from 5 to 18 years (EFICA), as well as to report the results of a comparison between children with ADHD and their peers without it. Thus, we conducted a population-based cross-sectional study relying on a sample composed of 3,284 typical children and adolescents accessed to study the psychometric properties of the parents' inventory (EFICA-P) and the teacher's inventory (EFICA-T) within a Structural Equation Modeling framework (SEM). Exploratory and confirmatory analyses were fitted, as well as the Cronbach's alpha and the McDonald's omega reliability indices. The known-groups method was carried out by independent Welch *t*-tests between untreated ADHD children and their peers. We concluded that the parents' inventory is composed of three dimensions (Cool Index 1, Cool Index 2, and Hot Index): $\chi^2(1,649) = 4,607.852$ $p < .01$, CFI = .965, TLI .963, RMSEA = .053, whereas the teachers' inventory is composed of two dimensions (Cool Index and Hot Index): $\chi^2(1,273) = 5,158.240$, $p < .01$, CFI = .991, TLI = .991, RMSEA = .077. The internal consistency of both inventories was $>.9$. Significant differences between the ADHD groups were found in all domains accessed. These findings indicate that both inventories have a high degree of validity regarding their internal structures, as well as supporting their clinical utility.

KEYWORDS

Psychometrics; executive functions; ADHD; child development

Introduction

Psychology in general and psychometrics in particular deals with latent process (factors, abilities), a class of variables not directly observed from the data, but supposed to exercise a causal influence to the observed behavior (Sijtsma, 2006). Analogous to other scientific disciplines, the theories within psychological science aim to enable one to explore, systematize, and predict observational experience by constructing models which mimic the phenomena of interest (Figueroa, 2016). On the other hand, there has been an increasing awareness that several theories suffer from a contradictory, incomplete, insufficient, or even a lack of evidence, even when considering theories as a fundamental piece of any scientific discipline within the psychological field, especially when the goal encompasses the univocal/specific definition of the objects intended to study (Boring, 1953; Muthukrishna & Henrich, 2019).

This statement holds true for many psychological processes, including in this group Executive Functions (EF). EF are elusive to define, therefore allowing the existence of different conceptualizations. This consequently nurtures the development of dissimilar standardized measures with

multiple results which are not entirely related or convergent (Barkley, 2001; Sijtsma, 2013). Broad definitions are often reported in the specific literature, and when one defines EF as *a broad cluster of frontal-lobe dependent high order cognitive and emotional abilities necessary for regulating and managing ongoing and goal-directed behaviors*, despite being an umbrella term, this explanation offers an important link which brings together fields such as Psychology, Psychiatry, and Neurology (Best & Miller, 2010; Blair & Raver, 2012; Diamond & Lee, 2011; Poon, 2018; Prencipe et al., 2011).

However, the accommodation of distinct theoretical roots does not always fit as expected. From a historical and academic perspective, EF as a concept comes with the tough mission of filling out or providing answers to some gaps on how specific cognitive processes are controlled and coordinated during the performance of complex cognitive tasks (Miyake et al., 2000). At least three broad questions emerge within the development of EF as a concept, and they can be further detailed. First, (i) which psychological processes comprise EF; second, (ii) how can these processes be grouped differently, and (iii) what are the roles of emotional and motivational aspects in EF? (Barkley, 2012; Gioia, Isquith, Retzlaff, & Espy, 2003; Goldstein & Brooks, 2013).

The processes which are involved in EF have undergone some metamorphoses throughout the years (Jurado & Rosselli, 2007). The first description was made by Baddeley and Hitch in 1974 within a framework in which the aim was to check whether the hypothesis of a “central executive” component was reasonable for a working memory model (Baddeley & Hitch, 1974). Next, in 1982, Lezak proposed to cluster EF into four major dimensions of human behavior which deal with “how” or whether behavior is expressed, in contrast to cognitive functions concerned with “what and how much knowledge, skill, and intellectual equipment a person may possess” (Lezak, 1982, p. 282). According to Lezak, these dimensions were capacities: (1) necessary for formulating goals; (2) involved in planning; (3) having to do with carrying out plans to reach the goals; and (4) for performing these activities effectively.

The Supervisory Attentional System (SAS) was proposed in virtually in the same period by Norman and Shallice (1986), wherein the authors explored the role of attention in behavior control. The interaction of all conceptualizations led to a contemporary and almost well-accepted EF core composed of inhibitory control, working memory, and cognitive flexibility (Diamond, 2013, Miyake et al., 2000). Despite this core, the attempt to explore the psychological processes related to EF is still appealing nowadays, and has originated a multifaceted mental concept, including more than 30 different components assumed to be interrelated (Barkley, 2001).

Within this context, a new theoretical EF framework has been recently proposed in the literature aiming at grouping EF processes, as represented by cool and hot EF (Peterson & Welsh, 2014; Poon, 2018; Prencipe et al., 2011; Schreiber, Possin, Girard, & Rey-Casserly, 2014). Cool EF are mainly composed of goal-directed and future-oriented behaviors. It comprises skills such as inhibition, flexibility, planning, working memory, and monitoring which are required/or emerge under analytical testing conditions, relatively decontextualized, and low- to non-emotional (Peterson & Welsh, 2014). On the other hand, hot EF is also related to goal-directed and future-oriented cognitive skills, but the main difference is the tension between immediate gratification and long-term reward, which elicits an interplay with emotion and motivation (Poon, 2018; Prencipe et al., 2011). Furthermore, other complex processes have been supplemented in addition to the classic tripod EF components, such as planning, organizing, attention, initiation, self-regulation, and self-monitoring (Gioia et al., 2003; Goldstein & Naglieri, 2014).

The role of emotion and motivation in conceptualizing EF has been neglected in the past, despite the evidence coming from animal models and human frontal lobe lesions which have been well-described in the seminal works of Luria, Damasio, and others (Damasio, 1994; Fuster, 1973; Spence, 1995). However, evidence coming from neurophysiological, neuropsychological, and clinical studies has changed the traditional concept of EF in the last two decades from purely cognitive skills which are elicited under relatively abstract, decontextualized, and non-affective

conditions to a broader and more comprehensive construct (Peterson & Welsh, 2014). The importance of emotion regulation on EF performance in different contexts and demands has led to consensus in the literature (Bechara, 2004; Zelazo, 2007), also overcoming a previous understanding that EF could be represented as a unidimensional structure (Gioia et al., 2003; Goldstein & Naglieri, 2014). In this same regard and according to Luria’s postulate and current evidence, the optimal interaction between neurobiological processes and environmental stimuli involving cultural, educational, and social factors is responsible for the cortical efficiency to process cognitive skills such as language, attention, memory, intelligence, and EF (Luria, 1966; Barkley, 2012; Gioia et al., 2003; Goldstein & Brooks, 2013). Concurrently with individual biological factors, culture modulates patterns of human knowledge development, belief system, and consequent behaviors. The impact of cultural and educational aspects on executive functioning has been shown in studies comparing children and adults from Western and Eastern countries (Goldstein & Brooks, 2013; Lan, Legare, Ponitz, Li, & Morrison, 2011; Luria, 1966).

Several reasons justify the importance of assessing EF in children and in adolescents, since EF are likely to have a significant impact on the development of everyday life skills. Since EF plays an essential role in cognitive development, its relationship with emotional regulation, and inhibitory control is regarded as part of the development of key outcomes of child development (Sun & Buys, 2012), as well for mental health (Han et al., 2016; Kertz, Belden, Tillman, & Luby, 2016; Shanmugan et al., 2016; White et al., 2017; Wilson, Andrews, Hogan, Wang, & Shum, 2018), socioemotional skills (Holmes, Kim-Spoon, & Deater-Deckard, 2016; Walk, Evers, Quante, & Hille, 2018), and school achievement (Cantin, Gnaedinger, Gallaway, Hesson-McInnis, & Hund, 2016; Oberer, Gashaj, & Roebbers, 2018; Viterbori, Usai, Traverso, & De Franchis, 2015).

The assessment of EF in children is challenging (Chevignard, Catroppa, Galvin, & Anderson, 2010; Miyake & Friedman, 2012), especially in terms of psychological instrumentation. Assuming Neuropsychology as the area interested to study the brain-behavior relationships, clinical observation and psychometrically defined behavioral examination go hand in hand (Barth et al., 2003). The development and use of neuropsychological testing has become a cornerstone for clinicians. These instruments are based on the assumption of an intrinsic relationship between cognition and behavior, and rely on the comparison between groups as well the performance obtained via single-subjects design. Thus, they enable to check how well various cognitive functions are operating, in addition to exploring which areas of the brain are functioning normally and which areas are having problems in helping to determine if a “person’s decision-making capacity may be impaired and why this might be happening” (Zasler, Katz & Zafonte, 2007, p. 1211).

Within assessment built on instrument-based strategies, test-based and questionnaire-based approach coexists and both play a pivotal role in clinical neuropsychology

(Batrancourt et al., 2010). With regard to test-based framework, performance-based neuropsychological tests comprise systematic procedures for their application, scoring, and interpretation of structured tasks. We call these instruments ecological measures (Woods et al., 2019) when these tools simulate or mimic daily life activities, or make this evaluation possible through behavioral observation with naturalistic stimuli. Finally, functional instruments refer to rating scales fulfilled by the parents, caregivers and/or teachers to evaluate child behaviors that mirror their executive functioning (Gillen, 2010; Lezak, Howieson, Bigler, & Tranel, 2012).

A lack of reliability, predictive value and ecological validity of performance-based tools have been argued in the literature (Barkley, 2012; Lezak et al., 2012). One of the most important limitations is the structured nature of the tasks, assigning the child's executive control, orienting, organizing, and self-monitoring to the examiner, and therefore masking the child's executive dysfunction (Barkley, 2012). Although some authors impute the lack of inter-observers and (inter-setting agreement) of functional scales and questionnaires as a cause of the limited validity and reliability of (functional assessment) (Chevignard et al., 2010), these instruments are able to provide useful clinical information on the discrepancy of child executive functioning in (different) contexts.

There are some functional rating scales to assess EF in children and adolescents (Goldstein & Naglieri, 2014), with the most cited being: the Behavior Rating Inventory of Executive Function (BRIEF) (Gioia et al., 2003), Barkley Deficits in Executive Functions Scale for Children and Adolescents (BDEFS-CA) (Barkley, 2012), Delis Rating of Executive Functions (D-REF) (Diamond & Lee, 2011), the Comprehensive Executive Function Inventory (CEFI) (Goldstein & Naglieri, 2014) and the Childhood Executive Function Inventory (CHEXI) (Thorell & Nyberg, 2008). Although reliable and valid instruments for functional assessment of EF in children and adolescents are mainly available in the United States, significant gaps still exist for Portuguese-speaking countries (such as Brazil, Portugal, and Mozambique), and communities within specific states in the United States.

First, the strong influence of cultural and socioeconomic aspects on EF requires enormous efforts to develop new instruments or adapt and validate the available ones across different languages and diverse realities, particularly in low and middle-income countries (Willoughby, Piper, Kwayumba, & McCune, 2019). Secondly, from the psychometric point of view, a very different methodology has been used to evaluate the construction, standardization, reliability, and validity of the available instruments, thereby making it difficult to compare them (Goldstein & Naglieri, 2014). Thirdly, diverse approaches are applied in the standardized samples used to create norms, for example, some authors only provide percentile scores (BDEFS-CA), others provide T-scores (BRIEF) or mean of 100 and standard deviation of 15 (CEFI) (Goldstein & Naglieri, 2014).

To the best of our knowledge, there is no specifically developed, validated and standardized functional instrument to evaluate EF in Brazilian children and adolescents. Of the

above-mentioned instruments, the BRIEF (Carim, Miranda, & Bueno, 2012) and the CHEXI (Trevisan, Dias, Berberian, & Seabra, 2017) have been translated, but only partially adapted to the Brazilian population. The data relied on convenience sampling with a lack of standardization or normative data. All analyses only achieved exploratory results, and the use of these tools nowadays additionally depends on payment due to copyrights, and consequently prevent their use, especially within the primary care area in developing countries.

Given the above, this study has two specific goals. First, by taking advantage of a large multistage populational study in which students were identified in the school system and the parents and teachers were directly interviewed, we aimed to develop an EF rating scale instrument for children and adolescents aged from 5 to 18 years. Second, we also sought to increase the clinical utility of this new tool by developing a percentile-based norm which can help professionals screen for the development of executive functions within children and adolescents.

Methods

Overview

The present study is a population-based cross-sectional study based on a multi-stage probability sample survey conducted in two cities located in the state of Minas Gerais, southeast region of Brazil. This is a stratified sampling method that takes into account the number of children in each place to adjust the final sample to the population information.

The first stage was conducted in 2013 in the city of São Sebastião do Paraíso, which covers an area with 64,980 inhabitants, 92.3% of them living in the urban area. According to the 2010 Brazilian demographic census, the Human Development Index (HDI) was 0.722, life expectancy 75.2 years, and the fecundity rate was 1.9 (IBdGeE, 2010c). The HDI is a summary measure for assessing long-term progress in three basic dimensions (a long and healthy life, access to knowledge, and a decent standard of living). Its scale ranges from 0 to 1, and the higher the score, the greater the development.

The second stage was carried out in 2016 in the city of Delfinópolis which has a total of 6,830 inhabitants, with 70.9% of them living in the urban area, HDI of 0.740, life expectancy of 77.3 years, and a fecundity rate of 2.3 (IBdGeE, 2010b).

The Federal Government in Brazil regulates the country's educational system through the Ministry of Education, which is equipped with public and private sector schools. While private schools are privately financed, public schools are funded by taxes and do not demand any direct payment. City governments coordinate the elementary schools and middle and high schools are managed by the state government. Because of the consortium characteristic of this study, access in São Sebastião do Paraíso was only permitted for elementary schools, whereas we could only evaluate children

in middle school in Delfinópolis. School is compulsory for all children between seven and 14 years of age.

Participants

Two different groups with responses of children and adolescent behaviors were composed for the analytical procedures of this research. The first group was mainly used to explore the internal structure of the instrument (its psychometric models), and therefore, we called it the validation group. The second group was called the clinical group, being useful to promote the clinical utility of the instruments and to demonstrate the relation with other variables and response processes.

An independent sample of children from two cities (São Sebastião do Paraíso, and Delfinópolis) were assembled to compose this first group. In São Sebastião, we aimed to assess all children from 5 to 12 years registered in any of the 12 public elementary schools, which are managed by the city government (1st to 5th grades in the American educational system). The target sample within these children was 3,658 participants, in which parental consents were obtained from 3,581 (97.9%), and 501 (14%) children were excluded due to the following criteria: diagnosis of Attention-Deficit/Hyperactivity Disorder (ADHD) ($n = 160$) or ASD ($n = 7$), visual impairment ($n = 94$), hearing impairment ($n = 47$), intellectual disability ($n = 37$), children using psychoactive drugs ($n = 40$), or with more than 10% of missing responses in the parent and/or teacher forms ($n = 116$). Thus, the final sample of this stage was 3,080 (84.2%) children.

In Delfinópolis, all children from 11 to 18 years registered in the only public school of the city with elementary and middle school (6th to 12th grades) ($n = 401$) were taken into account. Thus, the target sample in this city was 270 adolescents, in which parental consents were obtained from 253 (93.7%), and 49 (18.1%) were excluded due to the following criteria: diagnosis of ADHD ($n = 5$), visual impairment ($n = 8$), hearing impairment ($n = 3$), intellectual disability ($n = 4$), psychoactive drug usage ($n = 2$), or those with more than 10% of missing responses ($n = 27$). The final sample of this stage was 204 (75.6%) adolescents.

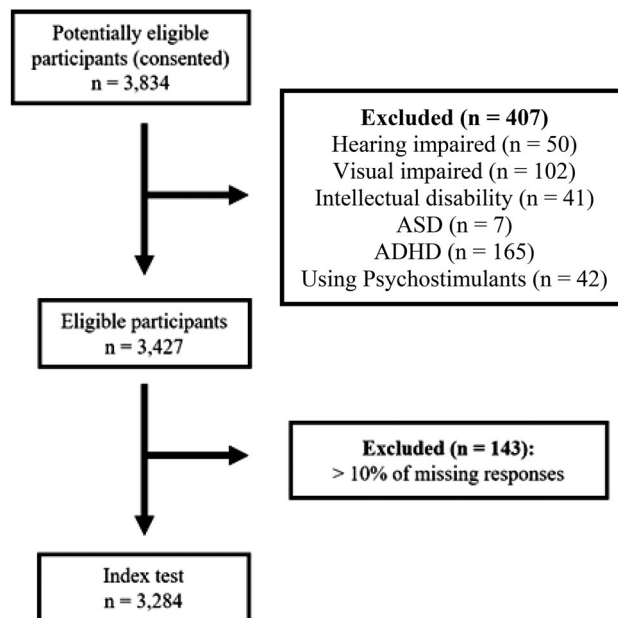
With that being said, the validation group was composed of 3,284 questionnaire responses from typical children and adolescents aged from 5 to 18 years of age ($Mean\ age = 8.2$, $SD = 2.0$), 50.1% boys, 11.1% living in rural area, 66% white, and 58.6% from middle income level (class C) (Table 1). As both scales were developed for parent-report, all data were reported by guardians or 173 school teachers (96% females).

Within this group, 79.8% ($n = 2,621$) of the parents' questionnaires were completed by the mother, 12.7% by the father, and 7.5% by another child's guardian. Children included in this sample were subjected to the following criteria: *I.* no visual, hearing, or intellectual impairment; *II.* no history of special education; *III.* no psychoactive drug usage; *IV.* Questionnaires with less than 10% of missing responses; and *V.* not fulfilling the last version of the Diagnostic and Statistical Manual of Mental Disorders (5th ed.; DSM-5) criteria for Attention Deficit Hyperactivity Disorder (ADHD)

Table 1. Target sample, consented, and final sample according to demographic features.

Age group	N	%
5 and 6 years old	1,017	31
7 and 8 years old	983	29.9
9 and 10 years old	807	24.6
11 and 12 years old	273	8.31
13 and 14 years old	88	2.68
15 and 16 years old	72	2.19
17 and 18 years old	44	1.34
$X^2(6) = 2,559, p < 0.01$		
Gender		
Female	1,638	49.9
Male	1,646	50.1
$X^2(7) = 0.02, p = 0.9$		
Race (self-reported)		
White	2,169	66
Nonwhite	863	26.3
Non-respondents	252	7.7
$X^2(2) = 1,752, p < 0.01$		
Area		
Urban	364	11.1
Rural	2,920	88.9
$X^2(1) = 1,989, p < 0.01$		
Income levels ^a		
A	42	1.3
B	1,027	31.3
C	1,925	58.6
D, E	290	8.8
$X^2(3) = 2,619, p < .01$		
House-head educational status		
Illiterate or incomplete middle school	1,602	48.8
Completed middle school but incomplete high school	651	19.8
Completed high school or more	956	29.1
Not determined	75	2.3
$X^2(3) = 40,540, p < .01$		
Total	3,284	100

^aIncome levels defined according to the Brazilian Economic Classification Criteria (A-ABdEd., 2015).



ASD: Autism Spectrum Disorder; ADHD: Attention Deficit Hyperactivity Disorder.

Figure 1. Flow chart shows sample population selection, eligible and excluded participants.

or Autism Spectrum Disorder (ASD). The number of children and adolescents of both stages selected in the different phases of the eligible process is shown in Figure 1.

The second group (clinical group) was formed by 165 children previously diagnosed with ADHD and not conducting any treatment at the time of the research. This group was formed of 126 (76%) of boys, with the mean age of 7.7 ± 2.16 years, and about 78% from families with low socioeconomic status. The data from these children were firstly collected through a semi-structured questionnaire based on the Swanson, Nolan and Pelham Questionnaire (SNAP-IV) Teacher 18-Item Rating Scale (Swanson, 2003), and filled out by their parents. We also asked the parents if their son/daughter has had a diagnosis of ADHD given by a specialized professional (physician or psychologist), and a specific question asking “Has your son/daughter taken Ritalin, Vyvanse, or Concerta in the past?”. These children were reevaluated via consultation with a medical doctor (pediatrician) or a psychologist with Specialization in neuropsychology, and their diagnosis was confirmed if the diagnostic criteria present in DSM-5 were present.

This study and the surveys received full approval from the Human Research Committee at the São José do Rio Preto School of Medicine (State of São Paulo). Written informed consent was discussed and obtained from the parents (or from the child’s guardian).

Measures

Executive function inventory for children and adolescents (EFICA)

Both EFICA-P (parents’ version) and EFICA-T (teachers’ version) comprises a standardized assessment tool developed to comprehensively assess dysfunctional behaviors related to EF. Their data are collected by individualized assessment only, in which it is expected that test users have knowledge of child development and are able to perform some mathematical computations in order to score and interpret the results adequately. These characteristics include EFICA-P and EFICA-T on the Level B test, following the traditional recommendations of the (American Psychiatric Association, 2013).

Parents and teachers are instructed to use a 0-2 rating scale indicating if each specific behavior occurs: never or almost never, (0), sometimes (1), or always/almost always (2). All items are summed to produce one final score in order to preserve its clinical utility, being inspired by other tools designed to assess child development, and higher scores indicate/suggest more problems (Gioia et al., 2003; Squires & Bricker, 2009; Squires, Bricker, Heo, & Twombly, 2001). Percentile cutoffs were used to aid clinicians to identify children needing further evaluation. Raw scores located at the 90th percentile are considered in the monitoring zone, and raw score located at the 95th percentile considers that the child could be at risk of developmental delay or disability in developing executive functions.

Similar to other measurement instruments within psychology and neuropsychology areas, the EFICA elaboration was based on a theoretical paradigm (Hot and Cold EF) and its relationship with other traditional standardized functional instruments used to assess EF in children and adolescents,

including BRIEF and CHEXI. Thus, the item development of the EFICA started with a comprehensive review of the literature and authors’ clinical and research experience based on best-practices recommendations (Roth, Isquith, & Gioia, 2014; Streiner, Norman, & Cairney, 2015). The first author has 30 years of clinical experience in evaluating and treating children and adolescents with executive dysfunction caused by a broad variety of causes.

All items were created to assess the traditional components of executive functioning, and the broad range of abilities dependent and/or related to executive functioning. Therefore, the content was created to assess the patient performance on real world behavior, i.e., daily routine tasks, as well to explore typical situation in which children should have an adequate emotional regulation or response with peers, and to plan, organize or modify their behavior to achieve the desired goals.

The Parents’ version (EFICA-P) encompassed behaviors especially performed at home, such as “Leaves the light on, door open or wet towels on top of the bed, even after being told several times”, “Explodes or gets angry when he/she is contradicted” and/or “Interrupts others, doesn’t know how to wait for his/her turn to talk.” The Teachers’ version (EFICA-T) had some particularities. As they are able to recognize child behaviors in the classroom, this version explored typical situations, asking if the child “gets up from their seat when they are supposed to be sitting”, “Loses their snack/sandwich, or their snack money and their homework” and/or “Leaves their belongings scattered everywhere they go throughout the classroom and the school”.

The initial pool of 180 items (95 items for the EFICA-P and 85 items for the EFICA-T) were reviewed and screened to identify redundancy, idiosyncrasy, and low specificity of EF domains according to the authors. As these are intensely debated in the literature, the fine differences and nuances which separate executive function deficits from ADHD and learning disorders can often be confusing, but since both clinical conditions are related, items with content which might overlap with other conditions were not summarily excluded in this part of the study (Barkley, 2001, 2012; Thorell & Wählstedt, 2006). This first step also intended to keep items immediately comprehensible to the target population, unambiguous, assertive, easy to use and culturally related (Streiner et al., 2015). A total of 63 items were removed, resulting in 65 items for the EFICA-P and 52 for the EFICA-T. The remaining 117 items were randomly combined in one pool and were submitted to two independent and licensed child neuropsychologists with more than 10 years of clinical practice. They were asked to identify the best fit EF domains for each item. When more than one domain was recognized, they were asked to rank them as primary or secondary domains.

The agreement between raters was measured by the Cohen’s kappa and reached a high value (k or $r \geq 0.75$) with .79 for EFICA-p ($p < .001$) and also .79 for EFICA-t ($p < .001$).

As recommended when developing new tools especially designed to assess child abilities, engaging small groups of

local key informants (i.e., parents and teachers working with young children) is an ideal way to collect information on the test content and procedures (Fernald, Prado, Kariger, & Raikes, 2017). Thus, the evidence based on test content (i.e., evidence associated with item content and item format or face and content validity) was also examined by two focus group discussions with the first author. They were not organized to generate the items, but with the objective for the research team to receive feedback on understanding the items from those with the same profile which EFICA was built to assess.

The first group was composed by six pairs of parents (EFICA-P) and the second one by eight teachers (EFICA-T) who were randomly selected from the 12 public elementary schools in Sao Sebastião, and from the only public school of the city with elementary and middle school (Delfinópolis) (see participants section for more details). A two-thirds majority consensus opinion was required. The recommended changes were mostly related to wording of the items and were implemented.

In summary, both questionnaires were developed aiming at capturing daily life behaviors related to executive functioning of children and adolescents aged from 5 to 18 years old, asking their parents and/or teachers to indicate if each specific behavior happens: never (or almost never), sometimes, or always (or almost always), on a three-point rating scale.

Procedure

One of the parents or the child's guardians was invited to attend a school meeting when the study was presented, and they completed the questionnaires while being supervised by the teachers and the first author. They also replied to a structured general questionnaire assessing sociodemographic characteristics, neuropsychomotor development, current and past health and medical history of the child, and other psychometric tools.

The teachers were trained by the first author on how to supervise the parents and also on how to fill out their questionnaires. Each school teacher in Brazil is responsible for approximately 20–30 children in their classroom, and therefore this was the number each teacher assessed. Demographic characteristics of teachers were only partially collected since there are strict laws protecting their information, and therefore it was not possible to nest responses at the teacher level.

Children from urban and rural areas were assessed as long as they were enrolled in the school system, which is required by law in Brazil.

Analytical plan

Different steps of the data analysis process were carried out in order to achieve the aims proposed in the present research. The dataset was initially explored aiming at finding the presence of inconsistencies and missing data. No outliers were excluded from the analysis. Multiple imputation procedure via Predictive Mean Matching (PMM) was used for

cases in which missing values were lower than or equal to 10%, whereas observations in which missing values were higher than this threshold were not included in the (Di Zio & Guarnera, 2010).

The data was described using plots, and descriptive statistics. The participants' characteristics were analyzed in terms of their age, gender, race, area of living, income level, and household head educational status. Race was divided into three categories: white, nonwhite, and non-respondents. The five income levels were defined according to the Brazilian Economic Classification Criteria, which are determined based on a combination of variables such as consumption of certain goods, and household characteristics (A-ABdEd, 2015).

A dataset composed of typical children was used ($n = 3,284$) to investigate the psychometric properties of both forms of EFICA (described in Participants section). The data was randomly divided into two subsets to execute the exploratory and confirmatory analyses in accordance with best practices in psychometrics literature. The exploratory analyses were conducted on one part of the sample, and the confirmatory analyses were performed on the second one (Costello & Osborne, 2011). Sample-splitting is generally considered a cross-validation method (Brown, 2015; Osborne, 2011).

The results of the Kaiser–Meyer–Olkin (KMO) and Bartlett's test were interpreted before performing an Exploratory Factor Analysis (EFA). A sample is deemed adequate if the value obtained by KMO is equal or greater than 0.6 in addition to a significant result on Bartlett's Test of Sphericity. This latter result indicates that the matrix is not an identity matrix, and is therefore suitable for EFA (Watson, 2017). Next, the EFA was performed to identify and determine the number and nature of the latent dimensions of the scales, as well as to reduce the number of variables being dealt with. All items were treated as categorical, and the estimator/extraction method used was the Weighted Least Squares Means and Variance adjusted (WLSMV), as it makes no distributional assumptions about the observed variables (Suh, 2015). Models with 2–5 oblique dimensions were estimated using a polychoric correlation. We note that CFA of categorical variables and Item Response Theory 2-PL models yield similar results.

The decision of how many factors to retain is a critical component of Factor Analysis procedures. In this study, the results of the eigenvalues scree-plot and the parallel analysis with 1,000 iterations were analyzed. The Confirmatory Factor Analysis (CFA) was then performed using the items from EFA which factor loadings were higher or equal to 0.3 (which represents about 9% of overlap in variance). Items with factor loadings in two or more factors simultaneously were preserved in its higher factor.

The CFA fit tends to be poorer than those achieved by EFA because constraints are implemented in CFA, whereas all indicators are loaded on all factors in EFA. Therefore, the CFA solution was considered appropriate if its Goodness of Fit (GoF) index met certain criteria and its interpretation was reasonable to the theoretical model. The

GoF criteria was met when the Comparative Fit Index (CFI) or the Tucker-Lewis Index (TLI) results were higher or equal than 0.9, and the Root Mean Square Error of Approximation (RMSEA) or the Standardized Root Mean Square Residual (SRMR) were less than 0.1 (Furr, 2011; Hu & Bentler, 1999). The reliability of the results was accessed via Cronbach's alpha, McDonald's omega, and the Average Variance Extracted within the confirmatory framework. A one-way ANOVA was carried out to check whether the difference between the results from children's guardians was significant.

Normative data and its respective tables were developed for both EFICA questionnaires based on the entire sample within a Classical Test Theory (CTT) framework specifically addressed to emphasize the clinical utility of the questionnaires and provide a useful resource to clinicians or primary care professionals. Thus, a linear model was fitted for EFICA-P and EFICA-T to check whether the gender, age, and the interaction between these two factors had an effect on the children's total results, being composed of raw scores of all the summed factors. A normative table was then created based on significant outcomes, being inspired by other questionnaires commonly used for assessing child development (Gioia et al., 2003; Squires & Bricker, 2009). Since the higher the score on the questionnaires, the greater the children's dysfunctional executive behaviors, the score related to 90th percentile was considered in the *monitoring zone*, and the score related to 95th percentile was considered that the child could be at risk of developmental delay or disability in developing executive functions, and so a further assessment is recommended to determine the need for intervention (Crawford & Garthwaite, 2009; Fonseca et al., 2006).

Finally, the known-groups method was used to further explore some features related to the validity evidence based on relations with other variables, as well as the clinical utility of both forms. Independent Welch *t*-tests between ADHD and non-ADHD groups were carried out in all investigated domains by both forms from a sample of 165 children previously diagnosed with ADHD according to the DSM-5 criteria, while Cohen's *d* was used as the measure of the effect size of the difference. The known-groups technique is considered a component of the traditional concept of validity studies and checks the extent to which a measure predicts different results between groups that are theoretically/previously known to differ (Hattie & Cooksey, 1984). The Welch test is more robust when compared to the traditional Student's *t*-test, and prevents type 1 error. All analyses were performed in (R Core Team, 2019) with *tidyverse* (Wickham, 2016), *lavaan* (Rosseel, 2012), *psych* (Revelle, 2015), and *ApaTables* (Stanley & Spence, 2018) packages.

Results

Reliability and validity

EFICA-P

As described in the methods section, the EFICA-P data were accessed to check its internal structure (mainly its reliability and validity). The internal consistency of all data obtained by Cronbach's alpha and the McDonald's omega coefficient

was .95 for both indices. All the items were positively correlated to the total score. However, the items 30 ("Can't pay attention to video games or computer games for a long time"), 33 ("Doesn't do a final check on his/her tasks or activities to look for possible mistakes"), 39 ("Doesn't know how to keep a secret"), 51 ("Recognizes they have a problem to solve"), and 64 ("Tries to solve a problem before asking someone for help") had an item-total correlation lower than 0.3.

The results obtained by the KMO and the Bartlett test were checked before performing the Exploratory Factor Analysis. The EFA was performed in half of the randomly divided data as the Bartlett test was significant at 5% probability level and the KMO was .97. A solution between 2 to 5 dimensions was explored and a solution with 3 dimensions had acceptable indexes in the confirmatory procedure: $\chi^2(1,649) = 4,607.852$ $p < .01$, CFI = .965 (Robust = 0.902), TLI .963 (Robust = 0.898), RMSEA = .053 (Robust = 0.044). The correlation between all dimensions were significant. Cool Index 1 and Hot Index correlated at .73, and Cool Index 1 with Cool index 2 correlated at .80. The Cool index 2 correlated at .70 with the Hot Index. Item properties are reported in Table 2.

The difference between the respondents/guardians total scores on the EFICA-P was evaluated and no statistically significant difference was found ($F(2, 3,214) = 2.12$, $p = .12$).

EFICA-T

The same process previously described was performed with the data gathered by the EFICA-T. The KMO was .97 and the Bartlett test was significant at 5% probability level. The Cronbach's alpha and McDonald's omega were .98. The high Cronbach's alpha results are a consequence of the majority of the respondents having checked the "never happens (or almost never)" category, as well as the high number of items in both questionnaires.

All items were positively correlated to the total score and the item-total correlation was higher than 0.3. The CFA revealed an adequate fit for the two-factor model: $\chi^2(1,273) = 5,158.240$, $p < .01$, CFI = .991 (Robust = .954), TLI = .991 (Robust = .952), RMSEA = .077 (Robust = .061). The correlation between both factors was .80 ($p < .01$). The results are reported in Table 3.

Normative data

EFICA-P and EFICA-T

A key feature of developing a new standardized measure is its normative data. Raw scores are often difficult to interpret and/or meaningless in the absence of further information. Therefore, normative data (and tables) provide information which enable drawing useful inferences about the person's relative performance. However, as EFICA-P and EFICA-T results were completed with data from boys and girls aged from 5 to 18 years old, testing the age and gender effects must be carried out to develop specific reference normative

Table 2. Items analysis, descriptive, and CFA results of EFICA-P (parents).

		Confirmatory Factor Analysis		Item description for the entire scale		
		λ Standardized (loadings)	λ Unstandardized	M	SD	r
Cool index 1						
8	Leaves the light on, door open or wet towels on top of the bed, even after being told several times.	0.715	1.082	1.08	0.71	0.53
20	Forgets to bring homework to school, even if they had done it.	0.699	1.058	0.32	0.54	0.44
17	Is lazy.	0.667	1.011	0.83	0.65	0.49
12	Messy, not able to keep his/her room, closet, or desk organized.	0.66	1	0.90	0.73	0.5
26	Leaves his/her clothes spread around the house.	0.656	0.994	1.00	0.74	0.51
48	When asked to go do something, he/she forgets what to do.	0.655	0.993	0.59	0.63	0.51
7	Makes careless mistakes	0.637	0.964	0.83	0.52	0.48
1	Leave everything until the last minute.	0.632	0.957	0.65	0.67	0.47
47	Loses their belongings or forgets where they put them.	0.617	0.935	0.83	0.68	0.53
10	It takes too long to do tasks such as bathing, changing clothes, and eating.	0.581	0.88	0.96	0.76	0.48
55	If you ask him/her to get three things from someplace, they will forget to bring one of them.	0.573	0.868	0.62	0.67	0.43
60	Has difficulty getting ready for school on time and ends up being late if someone isn't pushing them to get ready.	0.568	0.86	0.75	0.79	0.45
57	Their notebooks, agendas and binders are always disorganized.	0.529	0.802	0.72	0.74	0.41
6	Always arrives late to appointments or commitments.	0.497	0.752	0.27	0.5	0.35
<i>Cronbach's α =.89</i>						
<i>McDonald's ω =.73</i>						
<i>Average variance extracted =.39</i>						
Hot Index						
16	Instructions need to be repeated many times to be obeyed.	0.785	1.114	1.08	0.73	0.64
4	Acts without thinking, has difficulty controlling his/her behavior and reactions.	0.716	1.016	0.55	0.66	0.59
21	"Explodes"/gets angry when he/she is contradicted.	0.705	1	0.88	0.71	0.55
45	Loses his/her control easier than others.	0.703	0.997	0.53	0.66	0.55
19	Stubborn, everything has to be his/her way, inflexible.	0.697	0.989	0.80	0.72	0.56
23	Gets angry for no reason.	0.69	0.978	0.88	0.75	0.53
3	Overreacts excessively to little problems.	0.676	0.959	0.72	0.67	0.53
28	Mood changes for no reason.	0.658	0.934	0.70	0.68	0.53
25	Interrupts others, doesn't know how to wait for his/her turn to talk.	0.655	0.929	0.90	0.7	0.54
59	He/she has difficulty waiting their turn.	0.654	0.928	0.85	0.68	0.54
52	Resists accepting an alternative way to solve a problem, isn't flexible.	0.646	0.916	0.83	0.7	0.53
14	Excessively restless, never stops.	0.642	0.911	0.78	0.73	0.54
24	He/She becomes upset with changes in plans or new situations to them.	0.619	0.878	0.70	0.68	0.53
18	Is repetitive, speaks or asks the same things repeatedly.	0.615	0.873	0.84	0.74	0.53
50	Wants to have fun and get attention from others at all times.	0.606	0.86	0.77	0.71	0.51
53	Focuses on unimportant details and does not perceive "the big picture."	0.603	0.855	0.65	0.63	0.5
42	Doesn't tolerate frustrations, doesn't accept when something he/she wants goes wrong or doesn't work.	0.568	0.806	0.92	0.68	0.46
22	Talks too much.	0.565	0.802	1.17	0.71	0.44
32	Not able to wait for rewards, prefers to get 5 Reais today (5 Dollars) than wait a week to get 10 Reais (10 Dollars).	0.545	0.773	0.83	0.78	0.46
11	Reckless/fearless, gets involved in risky situations.	0.54	0.766	0.39	0.6	0.41
54	Gets frustrated and cries easily.	0.53	0.753	0.88	0.75	0.48
58	His/her moments of anger and crying are intense but end quickly.	0.444	0.63	0.98	0.74	0.34
<i>Cronbach's α =.93</i>						
<i>McDonald's ω =.83</i>						
<i>Average variance extracted = .40</i>						

(continued)

Table 2. Continued.

		Confirmatory Factor Analysis		Item description for the entire scale		
		λ Standardized (loadings)	λ Unstandardized	M	SD	r
Cool Index 2						
46	Misses deadlines to deliver homework or meet other commitments.	0.734	1.988	0.35	0.56	0.47
27	Even when knowing the answer, makes mistakes due to being careless, impulsive, or from a lack of attention.	0.712	1.928	0.78	0.61	0.5
37	Doesn't finish what they start and quits soon after encountering any difficulty.	0.679	1.838	0.69	0.68	0.56
44	Needs a "push" to start an activity, someone to take initiative for him/her.	0.675	1.829	0.77	0.71	0.55
5	When speaking, he/she doesn't finish one subject and moves on to another.	0.633	1.715	0.42	0.6	0.47
62	Has difficulty staying focused on homework and in the classroom.	0.613	1.662	0.70	0.67	0.51
31	Not good at solving problems without help from another person.	0.607	1.645	0.73	0.64	0.44
2	It takes time to understand things.	0.606	1.641	0.63	0.62	0.5
35	Doesn't perceive their own performance very well, leaves an exam thinking it went well, when in fact it didn't.	0.602	1.631	0.57	0.63	0.43
13	Not able to evaluate if a problem has been solved well.	0.579	1.569	0.69	0.6	0.45
38	Doesn't know how to act/overcome new situations or when confronting a problem.	0.563	1.525	0.81	0.63	0.42
34	Doesn't think before speaking or acting.	0.547	1.481	0.86	0.64	0.46
43	Needs help from an adult to finish an activity.	0.537	1.455	0.94	0.62	0.45
40	Doesn't have a good notion of the time necessary to perform a determined task.	0.517	1.399	0.84	0.67	0.41
36	Doesn't notice that his/her behavior is bothering other people.	0.5	1.355	0.77	0.68	0.41
9	It takes time to understand jokes.	0.499	1.352	0.56	0.61	0.41
56	If you make plans to pick them up at a certain time, they are never ready at the scheduled time.	0.458	1.241	0.60	0.7	0.33
41	Doesn't have creativity to solve day-to-day problems.	0.452	1.225	0.71	0.67	0.37
65	Understands things very literally, has difficulty understanding expressions with double/idiomatic meanings.	0.437	1.183	0.81	0.66	0.34
39	Doesn't know how to keep a secret.	0.415	1.125	0.88	0.74	0.28
29	Not able to pay attention to movies or cartoons for prolonged periods.	0.376	1.017	0.44	0.66	0.33
30	Can't pay attention to video games or computer games for a long time.	0.369	1	0.38	0.64	0.28
33	Doesn't do a final check on his/her tasks or activities to look for possible mistakes.	0.354	0.959	0.83	0.7	0.28
Cronbach's $\alpha = .90$						
McDonald's $\omega = .76$						
Average variance extracted = .31						
Cronbach's α (total) = .96						
McDonald's ω (total) = .90						
Average variance extracted (total) = .36						
χ^2 (1,649) = 4,607.852 $p < .01$, CFI = .965 (Robust = 0.902), TLI .963 (Robust = 0.898), RMSEA = .053 (Robust = 0.044)						

r means the correlation of the item with the scale composed of the remaining item.

table norms to assist in the classification or description of new examinees.

A linear model was fitted for EFICA-P results considering the results of all the summed items, and no interaction effect between age and gender ($F(1, 3,280) = 1.06, p = .302$), nor gender alone ($F(1, 3,280) = .52, p = .470$) was found. Contrasting these first results, a simple main effect revealed that children's age had a significant effect on the overall results, but with a small to null effect size ($F(1, 3,280) = 9.47, p = .002, \eta_p^2 \approx 0.01$).

On the other hand, the results of EFICA-T had a different pattern. Despite statistically null results for effects of age ($F(1, 3,280) = .46, p = .497$), and gender ($F(1, 3,280) = .46, p = .500$), an interaction effect was found between gender and age ($F(1, 3,280) = 17.35, p < .01, \eta_p^2 \approx .01$). When an interaction term is significant in a model, the interpretation of main effects is no longer the average effect but rather a conditional effect, and therefore this result reflected the development of the normative table for EFICA-T. Table 4 reports the results.

Table 3. Item analysis and CFA results – EFICA-T (Teachers).

		Confirmatory Factor Analysis		Item description for the entire scale		
		λ Standardized (loadings)	λ Unstandardized	M	SD	r
Cool index 1						
8	Has difficulty focusing on school tasks and work.	0.92	1.02	0.66	0.69	0.79
25	Doesn't have a good notion of the time required to perform a determined task.	0.91	1	0.56	0.66	0.73
23	Has difficulty completing tasks and activities.	0.91	1	0.54	0.67	0.75
19	Needs a "push" to start an activity, someone to take initiative for him/her.	0.9	0.99	0.47	0.67	0.75
48	Has difficulty maintaining concentration for prolonged periods.	0.89	0.98	0.67	0.7	0.78
14	Has difficulty performing tasks that require more than one step for their execution.	0.88	0.97	0.65	0.68	0.75
21	Does not finish homework and class activities.	0.88	0.97	0.5	0.65	0.72
30	Doesn't know how to find adequate resources to solve a problem.	0.88	0.97	0.61	0.64	0.71
17	Needs help and monitoring by an adult to keep doing a task or another activity	0.88	0.96	0.61	0.72	0.74
49	Doesn't finish an activity on time.	0.87	0.96	0.56	0.68	0.73
29	Starts their tasks at the last minute, puts things off/procrastinates.	0.87	0.96	0.33	0.59	0.7
36	Doesn't take notes well.	0.87	0.95	0.54	0.64	0.71
12	They are easily distracted by sounds and visual stimulations.	0.85	0.94	0.82	0.69	0.73
52	Doesn't understand simple instructions.	0.84	0.93	0.45	0.61	0.69
9	Doesn't use creativity to solve problems.	0.84	0.93	0.66	0.66	0.67
1	Doesn't perceive their own performance very well, leaves an exam thinking it went very well, when in fact it was very bad.	0.83	0.92	0.65	0.65	0.66
13	Forgets to bring homework to school, even if they had done it.	0.83	0.92	0.51	0.63	0.68
32	Doesn't try to solve a problem before asking the teacher for help.	0.83	0.92	0.5	0.64	0.68
33	Has difficulty perceiving their abilities and weaknesses.	0.82	0.91	0.58	0.64	0.71
3	Loses their snack/sandwich, or their snack money and their homework.	0.82	0.91	0.26	0.49	0.61
43	Has difficulty thinking about a different way to solve a problem.	0.82	0.9	0.64	0.66	0.66
39	Can't find what they need in the classroom or in their desk.	0.82	0.9	0.36	0.57	0.65
37	Has difficulty remembering things, even after only a few minutes.	0.81	0.89	0.41	0.59	0.66
46	Makes mistakes on tests even when knowing the answer for being careless, impulsive, or inattentive.	0.81	0.89	0.8	0.64	0.67
18	Focuses on unimportant details and does not perceive "the big picture."	0.79	0.87	0.45	0.62	0.68
38	Doesn't take initiative.	0.79	0.87	0.57	0.66	0.61
4	Doesn't bring their homework, agenda, or school materials.	0.77	0.85	0.53	0.66	0.6
6	Doesn't do a final check on their tasks or activities to look for possible mistakes.	0.76	0.84	0.78	0.68	0.62
7	Has good ideas, but can't organize them, plan them, or put them to paper.	0.75	0.83	0.69	0.64	0.6
47	Is sloppy on schoolwork and handwriting.	0.75	0.82	0.45	0.66	0.63
10	If you give them 3 things to go get, they forget one.	0.74	0.81	0.47	0.61	0.61
15	Doesn't ask for help when they need it.	0.69	0.76	0.55	0.64	0.56
Cronbach's $\alpha = .99$						
McDonald's $\omega = .94$						
Average variance extracted = .69						
Hot Index						
44	Leaves everything/everywhere messy and others need to clean up.	0.92	1.1	0.24	0.5	0.64
51	Is a disaster with their belongings, pens, pencils, erasers always falling on the floor.	0.92	1.1	0.29	0.56	0.68
28	Loses their temper more than their friends.	0.91	1.09	0.25	0.52	0.6
40	Leaves their belongings scattered everywhere they go throughout the classroom and the school.	0.9	1.08	0.26	0.52	0.65
35	Has problems when not supervised by an adult.	0.88	1.05	0.4	0.62	0.69
41	Doesn't think about the consequences of their actions.	0.87	1.05	0.5	0.63	0.67
11	Their desk is a mess.	0.86	1.04	0.34	0.59	0.65
22	Doesn't think before speaking or acting.	0.85	1.02	0.54	0.66	0.65
27	Gets up from their seat when they are supposed to be sitting.	0.85	1.02	0.44	0.64	0.61
34	Speaks and plays loudly and at inappropriate moments.	0.85	1.02	0.43	0.63	0.64

(continued)

Table 3. Continued.

	Confirmatory Factor Analysis		Item description for the entire scale		
	λ Standardized (loadings)	λ Unstandardized	M	SD	r
31 Resists or has difficulty accepting an alternative form of solving a problem with their friends, tasks, or school work, "it has to be done their way."	0.84	1.01	0.3	0.55	0.62
45 Has difficulty waiting their turn.	0.84	1.01	0.42	0.6	0.61
26 Interrupts others, doesn't wait for their turn to speak.	0.83	1	0.43	0.62	0.58
2 "Makes a mountain out of a molehill", overreacts excessively to little problems.	0.83	1	0.36	0.59	0.57
5 Gets upset over changing teachers or classes.	0.83	0.99	0.31	0.52	0.57
16 Has frequent mood swings for no reason.	0.83	0.99	0.34	0.58	0.57
50 Is repetitive, repeatedly speaks and asks the same thing.	0.82	0.99	0.25	0.51	0.6
42 Explodes/Becomes furious when they are contradicted.	0.81	0.97	0.37	0.6	0.55
24 Does not accept disappointments or reprimands.	0.81	0.97	0.48	0.64	0.59
20 Doesn't notice that their behavior is bothering other people.	0.79	0.95	0.57	0.67	0.6

Cronbach's $\alpha = .98$

McDonald's $\omega = .91$

Average variance extracted = .73

Cronbach's α (total) = .99

McDonald's ω (total) = .96

Average variance extracted (total) = .71

χ^2 (1,273) = 5,158.240, $p < .01$, CFI = .991 (Robust = .954), TLI = .991 (Robust = .952), RMSEA = .077 (Robust = .061)

Table 4. Normative table for EFICA-p and EFICA-t results.

Age group	EFICA-P (Parents' form)					EFICA-T (Teachers' form)					
	M	SD	P90	P95	N	Gender	M	SD	P90	P95	N
5 and 6 yo	44	18.9	69	77	1,017	Girls	21	20	50	61	498
						Boys	27	22	58	70	519
7 and 8 yo	45	19.7	71	79	983	Girls	22	20	51	60	492
						Boys	29	23	61	74	491
9 and 10 yo	43	19.6	70	76	807	Girls	20	19	46	55	414
						Boys	29	23	65	72	393
11 and 12 yo	43	20.5	71	77	273	Girls	21	18	42	55	128
						Boys	34	24	65	77	145
13 and 14 yo	42	21.9	72	81	88	Girls	20	23	56	65	46
						Boys	38	29	78	92	42
15 and 16 yo	40	19.3	67	72	72	Girls	20	22	52	53	35
						Boys	45	26	75	98	37
17 and 18 yo	31	18	57	60	44	Girls	24	20	52	58	25
						Boys	29	32	70	101	19
General	44	19.6	70	77	3,284		25	22	56	68	3,284

yo: years old.

We recommend using the suggested cutoff points as preliminary, incipient or exploratory for groups, whereas the sample size is fewer than 50 participants.

Known-Groups

When the different validity sources of a new tool are explored, an important method to access one of these sources is the "Known-Groups" method. The evidence of the test validity is seen within this procedure, as its scores significantly discriminate across groups which are theoretically known to differ. Consequently, this method potentiates the clinical utility of an instrument, which is not different for the scales studied in this research. Therefore, the clinical utility of both scales was then explored from these results.

The full sample ($n = 3,284$) was used to obtain information about its means and standard deviation to compare these results with those obtained from the ADHD sample ($n = 165$). All comparisons were carried out via an independent Welch Two Sample t -test. This test is robust for

variance heterogeneity and therefore robust in terms of Type I errors, and more preferable than the traditional Student's t -test.

The results from the parent form revealed a significant difference between the groups in all domains assessed. The Cool Index 1 was higher for ADHD participants ($M = 17.9$, $SD = 5.23$) than non-ADHD participants ($M = 10.3$, $SD = 5.50$; $p < .01$, $d = 1.4$ 95% CI [1.2, 1.5]), as well as for Hot Index (ADHD participants: $M = 30.00$, $SD = 8.1$; Non-ADHD participants: $M = 17.6$, $SD = 9.21$, $p < .01$, $d = 1.3$ 95% CI [1.20, 1.5]), and Cool Index 2 (ADHD participants: $M = 27.1$, $SD = 7.67$; Non-ADHD participants: $M = 15.7$, $SD = 7.62$, $p < .01$, $d = 1.5$ 95% CI [1.3, 1.7]). Taking into account that the higher the value of these indexes, the greater the difficulties presented by the child in these EF domains.

The same pattern of results was found regarding the teacher questionnaires. Cool Index results were higher in the ADHD group ($M = 44.55$, $SD = 13.10$) than the non-ADHD group ($M = 17.74$, $SD = 15.14$; $t(172.84) = 24.51$, $p < .01$,

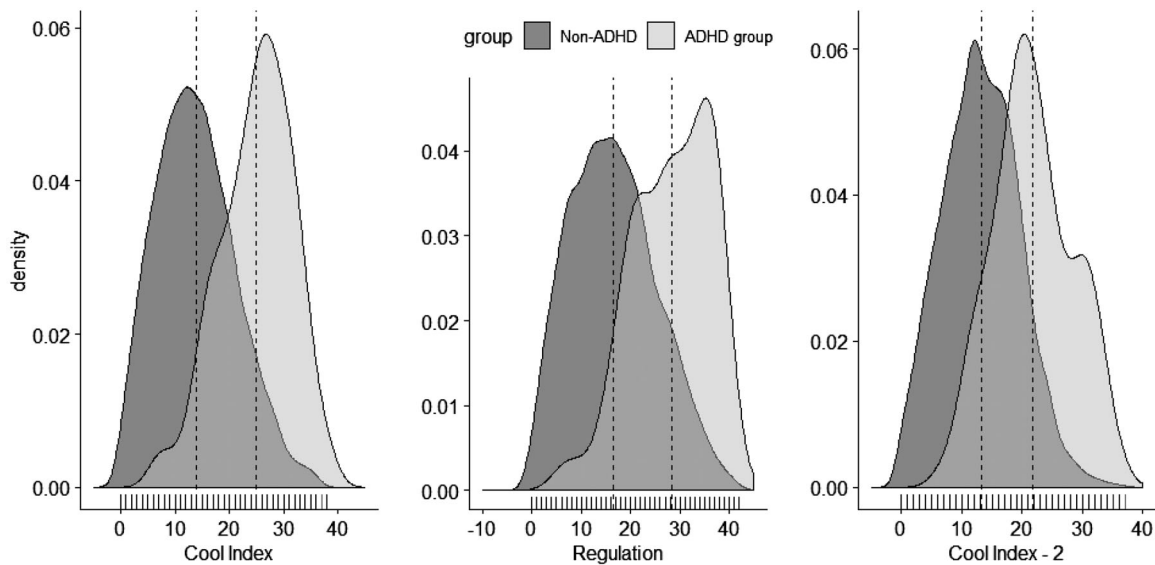


Figure 2. Known-Group results (EFICA-P).

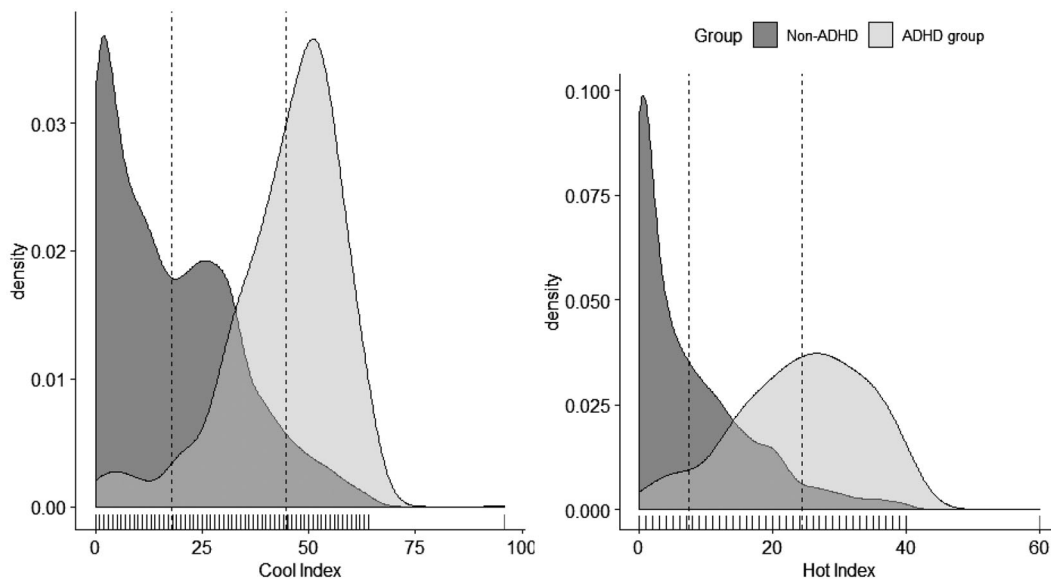


Figure 3. Known-Group results (EFICA-T).

$d = 1.78$ 95% CI [1.61, 1.95]), and the Hot Index results were also higher in the ADHD group ($M = 24.47$, $SD = 9.66$) than the non-ADHD group ($M = 7.48$, $SD = 8.54$; $t(165.9) = 21.49$, $p < .01$, $d = 1.98$ 95% CI [1.81, 2.14]). Figures 2 and 3 show the density differences between ADHD (light grey) vs. non-ADHD (dark grey) participants according to the EFICA-P and EFICA-T, respectively. As previously explained, the higher the result, the higher the risk for any of the EF problems.

Discussion

Previous studies indicate that EF plays an important role in cognition, emotions, and behavior, in addition to being correlated with key outcomes in child development, mental health, socio-emotional skills, and school achievement (Brock, Rimm-Kaufman, Nathanson, & Grimm, 2009;

Cantin et al., 2016; Han et al., 2016; Meltzer, 2007; Shanmugan et al., 2016). Although the importance of studying EF is evident, there is a scarcity of standardized instruments to perform a functional EF evaluation in Brazil and in other low and middle-income countries (Willoughby et al., 2019). To a lesser degree, this lack of a standardized tool is also evident in Portuguese-speaking communities in the United States and other countries.

Thus, in the present study we describe the developmental process and the psychometric properties of a new functional instrument to evaluate EF within a large population sample of Brazilian children and adolescents aged from 5 to 18 years-old. Normative tables were developed considering age groups and children gender, as these factors were significantly related to the outcomes. Percentiles were used as reference since they are easy to understand and have an intuitive interpretation as they report the percentage of observations lying below or equal to the raw score obtained by the child.

After successive analytical approaches, the present study shows robust evidence of the different sources of validity for both EFICA-P and EFICA-T. The internal structure of the scales was submitted to an empirical analysis within the framework of internal consistency reliability and categorical factor analysis, which is comparable to 2PL IRT model. Cronbach's alpha and the McDonald's omega coefficient met the required rules of thumb for their acceptance. The resulting coefficients range from 0 to 1 in providing this overall assessment of an instrument's reliability. A high Cronbach's alpha means that all of the items have high covariances, ensuring high internal consistency (Clark & Watson, 1995).

When exploring the dimensionality aspects of both scales, the solution obtained from the factor analysis processes encompasses a multidimensional structure. Therefore, it points out the complexity of EF, and allows growth to the conclusion of the multidimensional nature of EF instead of a unidimensional nature (Gioia et al., 2003; Goldstein & Naglieri, 2014). Although these results support a multidimensional structure, the high correlation between the different behaviors observed to execute different tasks related to EF also allows understanding that all EF are linked or mutually dependent. This result is in accordance with other conclusions reported in the literature (Castellanos, Kronenberger, & Pisoni, 2018; Gioia et al., 2003; Skogan et al., 2016). There is no restriction to model loadings during exploratory analyses, therefore they are obtained for all items among all factors. In contrast, the confirmatory analyses can be modeled using the exploratory results, but adding restrictions regarding which variables load into specific factors.

An interesting solution was obtained after analyzing the interface between the exploratory and confirmatory analysis. The EFICA-P was formed by three different factors, from which two factors are composed of distinct cool EF. The Cool Index 1 was composed of items pointing to goal-directed behavior, planning, organizing, initiating, attention, and working memory. Persistency and self-monitoring were mainly found in the Cool Index 2. The factor related to hot EF was composed of items such as inhibiting, flexibility, and emotional regulation. These results are convergent to a model to other findings in the literature (Diamond & Lee, 2011).

Some considerations should be taken into account for better understanding of the results of the two related, yet independent factors reflecting cool EF obtained by the EFICA-P, in contrast to just one factor to the same theoretical construct obtained by the EFICA-T. Although the results obtained by parents and teachers are in the same direction, when considering the set of instruments that inspired the development of the EFICA, but especially the BRIEF, parents tend to rate their children, both boys and girls, as having more problems on the BRIEF-P and BRIEF as compared to teachers. This is consistent with other literature on parent-teacher discrepancies (Offord et al., 1996). Jarratt et al. reported an average correlation of .58 (range

= .46–.72) in a sample of 40 children (Jarratt, Riccio, & Siekierski, 2005).

Discrepancies between parent and teacher ratings and differences in the inter-rater reliability of specific scales may partly reflect the consistency with which behavioral and emotional difficulties are expressed in disparate environments and demands. In addition, the results obtained in the normative study also accounted for these differences. For example, problems with disinhibition may be more readily observed across settings than some other aspects of executive dysfunction. Differences with respect to degrees of environmental structure and demand in home and school settings may also contribute to inconsistencies between raters.

The relationship between EF and gender also depends on the definition of EF, and consequently the procedures relied upon for its assessment. Some academic dispute is present in the field which partially happens because the evidence is not concise, whether there exists a difference only in certain behaviors (such as attention or working memory), or if the differences found are direct consequences of methodological choices, and not as a function of gender at all.

As an example, results from a recent meta-analysis about visual-spatial working memory concluded a significant but small advantage for males when compared to females (mean $d = 0.155$, 95% confidence interval = 0.087–0.223), with the expectation of memory for location, in which a female advantage emerged. Thus, these results support the idea of the existence of cognitive gender (Voyer, Voyer, & Saint-Aubin, 2017; Wang & Carr, 2014).

In contrast, another meta-analysis considering both humans and other animals found little support for significant gender differences in executive function (deemed as attention and focus, impulse control, decision making, and working memory). The authors considered that age of testing, as well as the specific testing modality were collapsing the gender effect (Grissom & Reyes, 2019).

In this study, we found that the age effect was significant for both parents and teachers, but the gender effect was only associated with the teachers' version of the EFICA. Some explanations can be made from this result and previous literature suggests a gender difference in externalizing behaviors, arguing that boys are more susceptible to attention problems than girls. Also, girls get higher teacher ratings for approaches to learning than boys when they are 6–7 years old (Berthelsen, Hayes, White, & Williams, 2017).

The relationship between all factors was positive, but the correlation between Cool Index 1 and 2 was higher than their association with Hot Index, supporting a current paradigm in the literature of a dual dimensionality of EF with a cool and hot components (Peterson & Welsh, 2014).

The second cool EF index distinguished by the parents is predominantly composed by two EF (self-monitoring and goal persistency), which can perhaps be better observed with a larger observation time, explaining why parents were able to identify them and teachers were not. The different perceptions about EF between parents and teachers in functional instruments are reported in the literature, such as different kinds of demand and time. In addition, teachers

spend less time with the child when compared to parents, and therefore it is expected they will have less knowledge about their behavior, abilities and difficulties (Germano, Brito, & Capellini, 2017; Jarratt et al., 2005).

Another result that must be dealt with is the statistical difference between the performance of the ADHD and non-ADHD groups. The capacity to discriminate between the groups expected to have different scores adds an important psychometric value to EFICA. However, clinical utility is also composed of sensitivity and specificity studies, which were not carried out in this study, and therefore, the use of EFICA to identify participants with (or in-risk of having) ADHD should be done with caution and in an exploratory context only (Lohr, 2002).

It is easy to recognize that items of both instruments address the common behaviors seen in children with ADHD, revealing the deep interplay between cognitive and behavioral demands, as well the reciprocal relation between attentional control and (Drigas and Karyotaki, 2017). ADHD has been associated with a broad range of negative outcomes for affected subjects on academic, occupational, and (Miller et al., 2013), as well as being associated with other mental and substance use disorders, criminality, and unemployment (Jensen & Steinhausen, 2015), causing a serious financial burden to families and society, which characterizes it as a major public health problem (Nasol, Lindly, Chavez, & Zuckerman, 2019; Reus-Smit & Snidal, 2009). The results found in this study are in line with the literature which reports the central role of executive dysfunction in ADHD diagnosis (Drigas and Karyotaki, 2016; Barkley, 2012; Schreiber et al., 2014).

The strengths of this study include the populational-based/robust sample size, the strong reliability and validity evidence, and the quite similar distribution of household-head education level compared to Brazilian standards (IBdGeE, 2010a). We emphasize the considerable influence of parents' educational status on the perception of executive functioning in children and adolescents (Goldstein & Naglieri, 2014). We also achieved some of the main criteria for normative tables. Both EFICA questionnaires have a clear definition of their objective, their items were retained after a continuous approach of studies of evidence of its validity, and the standardization of both test administration and scoring increase their clinical utility.

However, this study also presents some limitations. First, the studied sample was composed of children and adolescents enrolled in the Brazilian public-school system. Although this is a well-accepted method to investigate disease epidemiology in the pediatric population, children from higher income levels are less likely to study in public schools in Brazil. Adjustments would not perfectly address this topic, since the socioeconomic variables would be shifted toward the lower incomes. Therefore, a generalization of the results achieved in this study is limited to children. In this regard, because the participants were not native English-speaking children, this tool should be used for exploratory purposes or just as preliminary tools outside of this context. Despite its limitation, these results could be of significant

importance for the adaptation of the tool within the United States and other English-speaking countries/communities.

Secondly, our study lacks test-retest and criterion (concurrent) validity analyzes. However, we have to keep in mind that establishing concurrent validity is particularly difficult for an EF rating scale given the broad variability in the conceptualization, measurement approach, and psychometric characteristics (Goldstein & Naglieri, 2014). Thirdly, the neuropsychological literature for defining an ideal sample size for normative group from which the normative table will be developed varies from $n=5$ to $n=500$. Therefore, we suggest using the results from cell in which the sample size is equal than 50 as exploratory purposes only when sample size is fewer than 50 participants. However, some characteristics are often common. Finally, this study did not count on a random sample of children who do not have clinical diagnoses, which could distort the interpretation of the results for these children.

Numerous reasons are responsible for referring children for neuropsychological evaluations. Difficulties in achieving scholastic achievement, and problems with hyperactivity and attention are frequently dependent on executive functions and lead to significant limitations in the daily routines of children. Clinical neuropsychologists often rely on clinical and parental interviews to assess everyday functioning. Despite the utility of the method, its unstructured nature leaves it vulnerable to the pitfalls of clinical judgment. Therefore, clinicians might benefit from having a new screening test with adequate psychometric properties for their daily use. Thus, this study provides a new instrument to be used which can be helpful in the first or preliminary process of identifying children and adolescents with a risk of poor EF development, thereby enabling preventive and remedial interventions and better school support for those children who are facing difficulties.

The validity process of a neuropsychological measure is related to the inferences and interpretations obtained by the use of the test, but not the test itself. As a result, because it is inadequate to state that a test is valid (or invalid) given a set of fixed data, another study is currently ongoing aiming to check the differential item function of the EFICA-P and EFICA-T items and to build more refined normative data.

Acknowledgment

We would like to acknowledge and thank both Dr. Landeira-Fernandez (PUC-Rio) and Dr. Cristiano Fernandes (PUC-Rio) for the fruitful discussions about statistics, psychometrics, and measurement.

ORCID

Luis Anunciação  <http://orcid.org/0000-0001-5303-5782>

References

- A-ABdEd, P. (2015). *Brazilian economic classification criteria*. Retrieved from <http://www.abep.org/Servicos/Download.aspx?id=11>.
- American Psychiatric Association. (2013). *DSM-5. Diagnostic and statistical manual of mental disorders*.

- Baddeley, A. D., & Hitch, G. (1974). Working memory. In: G. H. Bower, G.H. (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47-89). New York: Academic Press. Retrieved from [http://dx.doi.org/10.1016/S0079-7421\(08\)60452-1](http://dx.doi.org/10.1016/S0079-7421(08)60452-1)
- Barkley, R. A. (2001). The executive functions and self-regulation: An evolutionary neuropsychological perspective. *Neuropsychology Review*, 11(1), 1–29.
- Barkley, R. A. (2012). *Executive functions what they are, how they work, and why they evolved*. New York, NY: The Guilford Press.
- Barth, J. T., Pliskin, N., Axelrod, B., Faust, D., Fisher, J., Harley, J. P., ... Silver, C. (2003). Introduction to the NAN 2001 Definition of a Clinical Neuropsychologist NAN Policy and Planning Committee. *Archives of Clinical Neuropsychology*, 18, 551–555. doi:10.1093/arclin/18.5.551
- Batrancourt, B., Dojat, M., Gibaud, B., Kassel, G. (2010). A core ontology of instruments used for neurological, behavioral and cognitive assessments. International conference on formal ontology in information systems, May 2010, Toronto, Canada. pp. 185–198, [ff10.3233/978-1-60750-535-8-185ff](https://doi.org/10.1007/978-1-60750-535-8-185ff). [ffinsem-00519801f](https://doi.org/10.1007/978-1-60750-535-8-185ff)
- Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition*, 55(1), 30–40. doi:10.1016/j.bandc.2003.04.001
- Berthelsen, D., Hayes, N., White, S. L. J., & Williams, K. E. (2017). Executive function in adolescence: Associations with child and family risk factors and self-regulation in early childhood. *Frontiers in Psychology*, 8. doi:10.3389/fpsyg.2017.00903
- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, 81(6), 1641–1660. doi:10.1111/j.1467-8624.2010.01499.x
- Blair, C., & Raver, C. C. (2012). Child development in the context of adversity: Experiential canalization of brain and behavior. *American Psychologist*, 67(4), 309–318. doi:10.1037/a0027493
- Boring, E. G. (1953). The role of theory in experimental psychology. *The American Journal of Psychology*, 66(2), 169. doi:10.2307/1418724
- Brock, L. L., Rimm-Kaufman, S. E., Nathanson, L., & Grimm, K. J. (2009). The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. *Early Childhood Research Quarterly*, 24(3), 337–349. doi:10.1016/j.ecresq.2009.06.001
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). New York, NY: Guilford Publications.
- Cantin, R. H., Gnaedinger, E. K., Gallaway, K. C., Hesson-McInnis, M. S., & Hund, A. M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *Journal of Experimental Child Psychology*, 146, 66–78. doi:10.1016/j.jecp.2016.01.014
- Carim, D. D. B., Miranda, M. C., & Bueno, O. F. A. (2012). *Tradução e adaptação para o português do behavior rating inventory of executive function – BRIEF translation and adaptation into portuguese of the behavior*. *Psicologia: Reflexão e Crítica*.
- Castellanos, I., Kronenberger, W. G., & Pisoni, D. B. (2018). Questionnaire-based assessment of executive functioning: Psychometrics. *Applied Neuropsychology: Child*, 7(2), 93–109. doi:10.1080/21622965.2016.1248557
- Chevignard, M. P., Catroppa, C., Galvin, J., & Anderson, V. (2010). Development and evaluation of an ecological task to assess executive functioning post childhood TBI: The children's cooking task. *Brain Impairment*, 11(2), 125–143. doi:10.1375/brim.11.2.125
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319. doi:10.1037/1040-3590.7.3.309
- Costello, A., & Osborne, J. (2011). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10, 1–9.
- Crawford, J. R., & Garthwaite, P. H. (2009). Percentiles please: The case for expressing neuropsychological test scores and accompanying confidence limits as percentile ranks. *The Clinical Neuropsychologist*, 23(2), 193–204. doi:10.1080/13854040801968450
- Damasio, A. R. (1994). *Descartes' error: Emotion, rationality and the human brain*. New York, NY: Putnam.
- Di Zio, M., & Guarnera, U. (2010). A multiple imputation approach to deal with the unity measure error. *Statistical Methods & Applications*, 19(3), 431–444. doi:10.1007/s10260-010-0132-1.
- Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, 333(6045), 959–964. doi:10.1126/science.1204529
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135–168. doi:10.1146/annurev-psych-113011-143750
- Drigas, A., & Karyotaki, M. (2017). Attentional Control and other Executive Functions. *International Journal of Emerging Technologies in Learning (Ijet)*, 12(03), 219 doi:10.3991/ijet.v12i03.6587.
- Drigas, A., & Karyotaki, M. (2016). Online and other ICT-based Assessment Tools for Problem-solving Skills. *International Journal of Emerging Technologies in Learning (Ijet)*, 11(04), 56 doi:10.3991/ijet.v11i04.5339.
- Fernald, L. C. H., Prado, E., Kariger, P., & Raikes, A. (2017). A toolkit for measuring early childhood development in low and middle income countries (English). Washington, D.C.: World Bank Group. Retrieved from <http://documents.worldbank.org/curated/en/384681513101293811/A-toolkit-for-measuring-early-childhood-development-in-low-and-middle-income-countries>
- Figuroa, A. (2016). The Alpha-Beta Method in Economics. In: *Rules for Scientific Research in Economics*. Cham: Palgrave Macmillan. Retrieved from https://doi.org/10.1007/978-3-319-30542-4_4
- Fonseca, A. C. C. F., Fonseca, M. T. M., Rodrigues, M. E. S. M., Lasmar, L. M. L. B. F., & Camargos, P. A. M. (2006). Pico do fluxo expiratório no acompanhamento de crianças asmáticas. *Jornal De Pediatria*, 82(6), 465–469. doi:10.1590/S0021-75572006000800012.
- Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. London: SAGE Publications Ltd. doi:10.4135/9781446287866
- Fuster, J. M. (1973). Unit activity in prefrontal cortex during delayed-response performance: Neuronal correlates of transient memory. *Journal of Neurophysiology*, 36(1), 61–78. doi:10.1152/jn.1973.36.1.61
- Germano, G. D., Brito, L. B., & Capellini, S. A. (2017). The opinion of parents and teachers of students with learning disorders regarding executive function skills. *Revista Cefac*, 19(5), 674–682. doi:10.1590/1982-0216201719510817
- Gillen, R. W. (2010). Neuropsychology of everyday functioning. *Archives of Clinical Neuropsychology*, 25(7), 685–686. doi:10.1093/arclin/acq055
- Gioia, G. A., Isquith, P. K., Retzlaff, P. D., & Espy, K. A. (2003). Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. *Child Neuropsychology*, 8(4), 249–257. doi:10.1076/chin.8.4.249.13513.
- Goldstein, S., & Brooks, R. B. (2013). *Handbook of resilience in children*. In S. Goldstein & R. B. Brooks (Eds.), Boston, MA: Springer US. doi:10.1007/978-1-4614-3661-4
- Goldstein, S., & Naglieri, J. A. (2014). *Handbook of executive functioning*. New York, NY: Springer-Verlag. doi:10.1007/978-1-4614-8106-5
- Grisson, N. M., & Reyes, T. M. (2019). Let's call the whole thing off: Evaluating gender and sex differences in executive function. *Neuropsychopharmacology*, 44(1), 86–96. doi:10.1038/s41386-018-0179-5
- Han, G., Helm, J., Iucha, C., Zahn-Waxler, C., Hastings, P. D., & Klimes-Dougan, B. (2016). Are executive functioning deficits concurrently and predictively associated with depressive and anxiety symptoms in adolescents? *Journal of Clinical Child & Adolescent Psychology*, 45(1), 44–58. doi:10.1080/15374416.2015.1041592
- Holmes, C. J., Kim-Spoon, J., & Deater-Deckard, K. (2016). Linking executive function and peer problems from early childhood through middle adolescence. *Journal of Abnormal Child Psychology*, 44(1), 31–42. doi:10.1007/s10802-015-0044-5
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. doi:10.1080/10705519909540118

- IbGe. (2010a). *Estatística da População Brasileira*. Retrieved from <https://www.ibge.gov.br/estatisticas-novportal/sociais/populacao/9662-censo-demografico-2010.html?edicao=9753&t=resultados>.
- IbGe. (2010b). *Estatística da População Brasileira - Delfinópolis*. Retrieved from <https://cidades.ibge.gov.br/brasil/mg/delfinopolis/panorama>.
- IbGe. (2010c). *Estatísticas da População Brasileira - São Sebastião do Paraíso*. Retrieved from <https://cidades.ibge.gov.br/brasil/mg/sao-sebastiao-do-paraíso/panorama>.
- Jarratt, K. P., Riccio, C. A., & Siekierski, B. M. (2005). Assessment of Attention Deficit Hyperactivity Disorder (ADHD) Using the BASC and BRIEF. *Applied Neuropsychology*, 12(2), 83–93. doi:10.1207/s15324826an1202_4
- Jensen, C. M., & Steinhausen, H.-C. (2015). Comorbid mental disorders in children and adolescents with attention-deficit/hyperactivity disorder in a large nationwide study. *ADHD Attention Deficit and Hyperactivity Disorders*, 7(1), 27–38. doi:10.1007/s12402-014-0142-1
- Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, 17(3), 213–233. doi:10.1007/s11065-007-9040-z
- Kertz, S. J., Belden, A. C., Tillman, R., & Luby, J. (2016). Cognitive control deficits in shifting and inhibition in preschool age children are associated with increased depression and anxiety over 7.5 years of development. *Journal of Abnormal Child Psychology*, 44(6), 1185–1196. doi:10.1007/s10802-015-0101-0
- Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, 108(3), 677–692. doi:10.1016/j.jecp.2010.11.001.
- Lezak, M. D. (1982). The problem of assessing executive functions. *International Journal of Psychology*, 17(1–4), 281–297. doi:10.1080/00207598208247445
- Lezak, Muriel Deutsch, Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York, NY: Oxford University Press.
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York: Oxford University Press.
- Lohr, K. N. (2002). Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of life research*, 11(3), 193–205. doi: 10.1023/a:1015291021312
- Luria, A. (1966). *The higher cortical functions in man*. Anuario De Psicologia Jurídica.
- Luria, A. R. (1966). *Higher cortical functions in man*. New York: NY: Basic Books.
- Marco Di Zio & Ugo Guarnera, 2008. “A multiple imputation method for non-Gaussian data,” *Metron - International Journal of Statistics*, LXVI(1), 75–90.
- Meltzer, L. (Ed.). (2007). *Executive function in education: From theory to practice*. Guilford Press.
- Miller, A. C., Keenan, J. M., Betjemann, R. S., Willcutt, E. G., Pennington, B. F., & Olson, R. K. (2013). Reading Comprehension in Children with ADHD: Cognitive Underpinnings of the Centrality Deficit. *Journal of Abnormal Child Psychology*, 41(3), 473–483. doi: 10.1007/s10802-012-9686-8.
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions. *Current Directions in Psychological Science*, 21(1), 8–14. doi:10.1177/0963721411429458
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. doi:10.1006/cogp.1999.0734
- Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour*, 3(3), 221–229. doi:10.1038/s41562-018-0522-1.
- Nasol, E., Lindly, O. J., Chavez, A. E., & Zuckerman, K. E. (2019). Unmet need and financial impact disparities for US children with ADHD. *Academic Pediatrics*, 19(3), 315–324. doi:10.1016/j.acap.2018.09.001
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour (Revised reprint of Norman and Shallice (1980)). *Consciousness and Self-Regulation: Advances in Research and Theory*. New York: Plenum Press.
- Oberer, N., Gashaj, V., & Roebbers, C. M. (2018). Executive functions, visual-motor coordination, physical fitness and academic achievement: Longitudinal relations in typically developing children. *Human Movement Science*, 58, 69–79. doi:10.1016/j.humov.2018.01.003
- Offord, D. R., Boyle, M. H., Racine, Y., Szatmari, P., Fleming, J. E., Sanford, M., & Lipman, E. L. (1996). Integrating assessment data from multiple informants. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(8), 1078–1085. doi:10.1097/00004583-199608000-00019
- Osborne, J. (2011). *Best practices in quantitative methods*. Thousand Oaks, CA: Sage Publications. doi:10.4135/9781412995627
- Peterson, E., & Welsh, M. C. (2014). The development of hot and cool executive functions in childhood and adolescence: Are we getting warmer?. In *Handbook of executive functioning* (pp. 45–65). New York, NY: Springer New York. doi:10.1007/978-1-4614-8106-5_4
- Poon, K. (2018). Hot and cool executive functions in adolescence: Development and contributions to important developmental outcomes. *Frontiers in Psychology*, 8(2311). doi:10.3389/fpsyg.2017.02311
- Prencipe, A., Kesek, A., Cohen, J., Lamm, C., Lewis, M. D., & Zelazo, P. D. (2011). Development of hot and cool executive function during the transition to adolescence. *Journal of Experimental Child Psychology*, 108(3), 621–637. doi:10.1016/j.jecp.2010.09.008
- R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org/>.
- Reus-Smit, C., & Snidal, D. (2009). *The Oxford handbook of international relations*. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199219322.001.0001
- Revelle, W. (2015). Package “psych” - Procedures for Psychological, Psychometric and Personality Research. R Package, 1–358. Retrieved from <http://personality-project.org/r/psych-manual.pdf>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 36. doi:10.18637/jss.v048.i02.
- Roth, R. M., Isquith, P. K., & Gioia, G. A. (2014). Assessment of executive functioning using the Behavior Rating Inventory of Executive Function (BRIEF). In S. Goldstein & J. A. Naglieri (Eds.), *Handbook of executive functioning* (p. 301–331). Springer Science + Business Media. Retrieved from https://doi.org/10.1007/978-1-4614-8106-5_18
- Schreiber, J. E., Possin, K. L., Girard, J. M., & Rey-Casserly, C. (2014). Executive function in children with attention deficit/hyperactivity disorder: The NIH EXAMINER battery. *Journal of the International Neuropsychological Society*, 20(1), 41–51. doi:10.1017/S1355617713001100
- Shanmugan, S., Wolf, D. H., Calkins, M. E., Moore, T. M., Ruparel, K., Hopson, R. D., ... Satterthwaite, T. D. (2016). Common and dissociable mechanisms of executive system dysfunction across psychiatric disorders in youth. *American Journal of Psychiatry*, 173(5), 517–526. doi:10.1176/appi.ajp.2015.15060725
- Sijtsma, K. (2006). Psychometrics in Psychological Research: Role Model or Partner in Science? *Psychometrika*, 71(3), 451–455. doi:10.1007/s11336-006-1497-9.
- Sijtsma, K. (2013). *Theory development as a precursor for test validity* (pp. 267–274). New York, NY: Springer.
- Skogan, A. H., Egeland, J., Zeiner, P., Øvergaard, K. R., Oerbeck, B., Reichborn-Kjennerud, T., & Aase, H. (2016). Factor structure of the Behavior Rating Inventory of Executive Functions (BRIEF-P) at age three years. *Child Neuropsychology*, 22(4), 472–492. doi:10.1080/09297049.2014.992401
- Spence, S. (1995). Descartes’ error: Emotion, reason and the human brain. *BMJ*, 310(6988), 1213–1213. doi:10.1136/bmj.310.6988.1213
- Squires, J., Bricker, D., Heo, K., & Twombly, E. (2001). Identification of social-emotional problems in young children using a parent-completed screening measure. *Early Childhood Research Quarterly*, 16(4), 405–419. doi:10.1016/S0885-2006(01)00115-6.

- Squires, J., & Bricker, D. (2009). *Ages and stages questionnaires. 3rd ed. (ASQ-3): A parent-completed child-monitoring system*. Baltimore, MD: Brooks Publishing Company.
- Stanley, D. J., & Spence, J. R. (2018). Reproducible tables in psychology using the apatables package. *Advances in Methods and Practices in Psychological Science*, 1(3), 415–431. doi:10.1177/2515245918773743
- Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use* (5th ed.). Oxford, UK: Oxford University Press. Retrieved from <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc11&NEWS=N&AN=2014-48611-000>.
- Suh, Y. (2015). The performance of maximum likelihood and weighted least square mean and variance adjusted estimators in testing differential item functioning with nonnormal trait distributions. *Structural Equation Modeling*, 22(4), 568–580. doi:10.1080/10705511.2014.937669
- Sun, J., & Buys, N. (2012). Early executive function deficit in preterm children and its association with neurodevelopmental disorders in childhood: A literature review. *International Journal of Adolescent Medicine and Health*, 24(4), 291–299. doi:10.1515/ijamh.2012.042
- Swanson, J. M. (2003) SNAP-IV Teacher and Parent Ratings Scale. In: F. Aykr (Ed.), *Therapist's Guide to Learning and Attention Disorders* (pp. 487–500). New York: Academic Press. Retrieved from <https://doi.org/10.1016/B978-012256430-7/50022-3>
- Thorell, L. B., & Nyberg, L. (2008). The Childhood Executive Functioning Inventory (CHEXI): A new rating instrument for parents and teachers. *Developmental Neuropsychology*, 33(4), 536–552. doi:10.1080/87565640802101516
- Thorell, L. B., & Wåhlstedt, 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. doi:10.1002/icd.475
- Trevisan, B. T., Dias, N. M., Berberian, A. D. A., & Seabra, A. G. (2017). Childhood executive functioning inventory: Adaptação e propriedades psicométricas da versão Brasileira. *Psico-USF*, 22(1), 63–74. doi:10.1590/1413-82712017220106
- Viterbori, P., Usai, M. C., Traverso, L., & De Franchis, V. (2015). How preschool executive functioning predicts several aspects of math achievement in Grades 1 and 3: A longitudinal study. *Journal of Experimental Child Psychology*, 140, 38–55. doi:10.1016/j.jecp.2015.06.014
- Voyer, D., Voyer, S. D., & Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: A meta-analysis. *Psychonomic Bulletin and Review*, 24(2), 307–334. doi:10.3758/s13423-016-1085-7
- Walk, L. M., Evers, W. F., Quante, S., & Hille, K. (2018). Evaluation of a teacher training program to enhance executive functions in preschool children. *PLOS One*, 13(5), e0197454. doi:10.1371/journal.pone.0197454
- Wang, L., & Carr, M. (2014). Working memory and strategy use contribute to gender differences in spatial ability. *Educational Psychologist*, 49(4), 261–282. doi:10.1080/00461520.2014.960568
- Watson, J. C. (2017). Establishing evidence for internal structure using exploratory factor analysis. *Measurement and Evaluation in Counseling and Development*, 50(4), 232–238. doi:10.1080/07481756.2017.1336931
- White, L. K., Moore, T. M., Calkins, M. E., Wolf, D. H., Satterthwaite, T. D., Leibenluft, E., ... Gur, R. E. (2017). An evaluation of the specificity of executive function impairment in developmental psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(11), 975–982.e3. doi:10.1016/j.jaac.2017.08.016
- Wickham, H. (2016). tidyverse: Easily Install and Load “Tidyverse” Packages. R package version 1.0.0. Retrieved from <https://cran.r-project.org/package=tidyverse>
- Willoughby, M. T., Piper, B., Kwayumba, D., & McCune, M. (2019). Measuring executive function skills in young children in Kenya. *Child Neuropsychology*, 25(4), 425–444. doi:10.1080/09297049.2018.1486395
- Wilson, J., Andrews, G., Hogan, C., Wang, S., & Shum, D. H. K. (2018). Executive function in middle childhood and the relationship with theory of mind. *Developmental Neuropsychology*, 43(3), 163–182. doi:10.1080/87565641.2018.1440296
- Woods, S. P., Kordovski, V. M., Tierney, S. M., & Babicz, M. A. (2019). The neuropsychological aspects of performance-based Internet navigation skills: A brief review of an emerging literature. *The Clinical Neuropsychologist*, 33(2), 305–326. doi:10.1080/13854046.2018.1503332.
- Zasler N., Katz D., Zafonte R., Eds. (2007). *Brain Injury Medicine: Principles and Practice*. New York, NY: Demos Medical Publishing.
- Zelazo Cunningham, W. P. D. (2007). Executive function: Mechanisms underlying emotion regulation. *Handbook of emotion regulation*. New York, NY: Guilford. doi:10.1177/1545968309354536