The Dutch version of the Behavior Rating Inventory of Executive Function-2 (BRIEF-2)

Huizinga, M.; Smidts, D.P.; Baeyens, D.; Kan, K.-J.

DOI
10.1027/2698-1866/a000038

Publication date
2023

Document Version
Final published version

Published in
Psychological Test Adaptation and Development

Original Article

The Dutch Version of the Behavior Rating Inventory of Executive Function-2 (BRIEF-2)

Mariëtte Huizinga[1,2], Diana P. Smidts[3], Dieter Baeyens[4], and Kees-Jan Kan[5]

[1]Department of Education and Family Studies, Vrije Universiteit Amsterdam, The Netherlands
[2]Research Institute LEARN!, Amsterdam, The Netherlands
[3]Kinderpsy Praktijk Hilversum, The Netherlands
[4]Parenting and Special Education, KU Leuven, Belgium
[5]Research Institute of Child Development and Education, University of Amsterdam, The Netherlands

Abstract: The BRIEF-2 (Gioia et al., 2015) is a widely used questionnaire to measure daily behavior related to executive function behaviors in the home and school environment of children between 5 and 18 years old. The current study was conducted to investigate the psychometric properties of the Dutch version of the BRIEF-2 in a representative Dutch-speaking norm sample. Using methods from classical test theory and network theory, we examined the reliability and validity of the BRIEF-2. The results indicated that the BRIEF-2 can be considered as a valid and reliable questionnaire that provides information on the role of executive function in the child's and adolescent's functioning in the home and school environment.

Keywords: executive function, BRIEF-2, development, questionnaire, children

The second edition of the Behavior Rating Inventory of Executive Function (BRIEF-2) is the first revision of the BRIEF questionnaire (Gioia et al., 2000). The BRIEF-2 is a questionnaire completed by parents and teachers of school-aged children or adolescents (5–18 years) and by adolescents themselves (11–18 years; self-report). With this questionnaire, daily behavior related to executive function can be assessed in the home and school environment. The BRIEF-2 is intended for a wide range of children or adolescents about whom there are concerns regarding self-regulation (Gioia et al., 2015).

Executive function[1] is generally defined as an umbrella term for various cognitive processes that subserve goal-directed behavior (Miller & Cohen, 2001; see also Luria, 1966; Shallice, 1982). Executive function is especially important in novel or demanding situations (Stuss, 1992), which require a rapid and flexible adjustment of behavior to the changing demands of the environment (Zelazo et al., 2003). The development of executive function is a protracted process, which extends into early adulthood (e.g., Diamond, 2013; Huizinga et al., 2006). Deviations in the development of executive function are often related to several symptoms and disorders, such as learning problems, anxiety, depression, trauma, and various developmental disorders and medical conditions (e.g., Hughes, 2011; Huizinga et al., 2018; Karr et al., 2018). Problems with executive function may be manifested in impulsive behavior, difficulties in planning ahead, and in adapting behavior to changing circumstances.

Over the past years, there has been substantial debate on the definition and measurement of executive function (Baggetta & Alexander, 2016; Karr et al., 2018). This debate centered on the unidimensionality (i.e., executive function comprises one mechanism without underlying factors) versus multidimensionality (i.e., executive function comprises multiple underlying factors) of executive function. In a recent systematic review, Karr et al. (2018) found some evidence for greater unidimensionality of executive function among child/adolescent samples and

[1] In the scientific literature, the label "executive function" is often used synonymously with "executive functions," referring to specific individual abilities, components, subcomponents, or processes or some combination of individual abilities (Baggetta et al., 2016). Here, we refer to "executive function" as an overall construct and to "executive functions" to label individual components.

https://doi.org/10.1027/2698-1866/a000038

both unity and diversity among adult samples. However, Karr et al. (2018) frequently observed low rates of model acceptance and selection. This was interpreted as an indication for possible bias toward publication of well-fitting but potentially nonreplicable models with underpowered samples. Thus, the debate on the conceptualization and consequent operationalization remains.

In 2015, the BRIEF-2 was published in the United States; in 2020, its Dutch variant was published in the Netherlands. The main goal of this paper was to examine the psychometric properties of the Dutch version of the BRIEF-2 in a representative Dutch-speaking norm sample.

BRIEF-2 Development

The idea for the BRIEF (Gioia et al., 2000) originated in the United States in 1994 when the authors observed that children's or adolescents' performance at executive function tasks often differed from reports by parents and teachers about their daily functioning, as well as from the authors' own clinical observations during assessments. The authors therefore developed a psychometrically well-researched measure of executive functions in children and adolescents which was easy to administer and assess and provided clinically useful information on widely accepted domains of executive function.

The original eight BRIEF scales (parent and teacher version) were based on a review of the available literature on executive functions, especially from studies focused on children and adolescents. Colleagues in clinical psychology, school psychology, and neuropsychology were asked about their conceptualizations of executive functions, and concise definitions for each scale were established. The authors collected possible questions based on clinical interviews with parents, teachers, and children or adolescents. Several procedures were used to reduce the original list of questions to the final scales: (1) mutual reviewer agreement among 12 (neuro)psychologists on how well the questions fit the different domains, (2) question properties such as frequent missing responses or poor distinction between clinical and normally developing groups, (3) iterative item-total analyses and exploratory factor analysis in large clinical data sets, (4) teacher assessment of item readability and clarity, and (5) expert review of the questions. This resulted in eight theoretically consistent scales with good empirical support. Principal factor analysis supported the merging of three scales (Inhibition, Shift, and Emotional Control) into a Behavioral Regulation Index and the remaining five scales (Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitoring) into a Metacognition Index (Gioia et al., 2001).

The BRIEF Self-Report (Guy et al., 2004) was later developed in a similar manner. The questions were derived from the BRIEF versions for parents and teachers and rewritten to be appropriate for adolescents. Additional questions were collected by the authors in clinical practice, and additional ones were added to ensure complete domain coverage. A similar analysis based on question characteristics, item-total analyses, factor analysis, and expert reviews led to the final scales. The 80 items formed eight scales, which were combined into two indices: Inhibition, Shift, Emotional Control, and Monitor comprised a Behavioral Regulation Index, and Working Memory, Plan/Organize, Organization of Materials, and Task Completion comprised a Metacognition Index.

In the Netherlands, the BRIEF (parent and teacher version) was first issued in 2009 (Huizinga & Smidts, 2011; Smidts & Huizinga, 2009); it was supplemented by the self-report in 2012 (Huizinga & Smidts, 2012).

Development and Standardization of the BRIEF-2

Since its publication in 2000, the BRIEF has experienced steady international growth in a wide range of application areas, including neuropsychology, clinical psychology, school psychology, education, medicine, and scientific research. With this growth, a wealth of knowledge about executive functions across the age spectrum has come to light. From this, several important factors emerged which necessitated a revision of the original BRIEF, for instance the general acceptance of a complex model of executive functions, characterized by behavioral, emotional, and cognitive regulation components (Grieve et al., 2014; Suchy, 2009), and the need for up-to-date norm data for reliable interpretation (Gioia et al., 2015).

The revision of the BRIEF did not involve a substitution or addition of items. Instead, the authors improved the BRIEF by taking four steps. First, they selectively pruned the scales to create a shorter instrument while maintaining the reliability and validity features of the original. Second, they increased the sensitivity to problems in executive functions by removing questions that did not contribute adequately to them. Third, they increased similarity between the three versions to make them easier to compare. Fourth, they collected an entirely new, representative norming sample.

The revision of the BRIEF began by reviewing all original questions from the BRIEF Parent and Teacher forms and Self-Report for clarity, syntactic consistency, and readability. When necessary, questions were clarified or made more specific. Subsequently, the content of the scales was reexamined to ensure that each scale is sufficiently diverse to cover the breadth and depth of each

domain of executive functions, based on current knowledge in practice and science. The items and scales were then iteratively examined within large, demographically diverse data sets (n > 2,000), in clinical and normally developing children and adolescents. Item characteristics were examined individually and with item-total analyses to determine their contribution to each scale. Scale-level exploratory factor analyses were used to clarify which items belonged to which scale (Gioia et al., 2015). Within this iterative process, the total item pool became smaller but maintained the strong support for content validity.

BRIEF-2 Clinical Scales

The following sections describe the content and interpretation of the BRIEF-2 individual clinical scales (each representing a domain of executive function) and indexes. Please see Tables 8–13 for the psychometric properties of the respective scales and indexes.

Inhibition

This scale measures the ability to resist or slow down an impulse and thereby also the ability to stop one's own behavior at the right time. Inhibition is a well-researched behavioral regulatory function that is seen as the core deficiency in attention-deficit/hyperactivity disorder (ADHD), primarily in hyperactive/impulsive subtypes (American Psychitric Association, 2013). It should be noted that weak inhibition is not specific to a particular disorder but is a general underlying problem in weak self-regulation (e.g., Barkley, 2012).

Shift

The Shift scale measures the ability to switch freely from one situation, activity, or dimension to another as circumstances demand. Key aspects of flexibility include the ability to make transitions, solve problems flexibly, and shift focus from one line of thinking or topic to another. Weak flexibility occurs in a variety of developmental disorders; the Shift scale may be the most elevated BRIEF scale in individuals with autism spectrum disorder (ASD; Hovik et al., 2017; Kenworthy et al., 2014).

Working Memory

The Working Memory scale measures the ability to process and hold information for short periods of time for the purpose of a task or activity. Working memory is essential for performing multistep activities, doing mental math, or following complex instructions. Weak working memory is observed in a variety of clinical populations involving executive function problems. Given the primary relationship between working memory and the diagnostic criteria for ADHD, the Working Memory scale may be useful in examining whether ADHD is present (McAuley et al., 2010; Toplak et al., 2009).

Self-Monitor

This scale measures the degree of awareness of the impact of one's own behavior on other people. It involves the ability to observe and evaluate one's own behavior as perceived by others. Although not formally included in the diagnostic criteria of developmental disabilities, inadequate awareness of one's own behavior is prominent in children or adolescents with, for example, ADHD, ASD, and traumatic brain injury (Gioia et al., 2015).

Task Behavior

Awareness of one's own behavior relates not only to social interactions (as with the Self-Monitor scale) but also to tasks and activities. The questions on this scale measure the degree of awareness of one's own actions in the context of performing tasks and activities. Task Behavior is not necessarily a stand-alone executive function, but for many children, it is a consequence of problems with different executive functions (Gioia et al., 2015).

Emotion Regulation

This scale focuses on the manifestation of executive functions within the emotional domain and a child's ability to modulate emotional responses. The link between executive functions such as inhibition or flexibility and emotion regulation is increasingly recognized (Beer et al., 2010; Gyurak et al., 2012; Perlman et al., 2014; Zelazo & Cunningham, 2007).

Plan/Organize

The Plan/Organize scale measures the child or adolescent's ability to oversee and direct current and future task demands. Although the constructs' planning and organizing are conceptually different, the observable behaviors are very similar. Analysis of the scale structure of the original BRIEF showed that the two scales were highly correlated, so they were merged into a single scale (Gioia et al., 2015). The planning component of this scale refers to the ability to anticipate future events, set goals, and choose appropriate steps in advance to carry out a task or activity. The organizing component of this scale refers to the ability to bring order to information and to recognize the main ideas or key concepts when learning or communicating information. The way information is strategically organized can play a crucial role in learning, remembering, and retrieving information. Difficulty with planning and organizing is often at the heart of self-regulation problems (Barkley, 2012).

Initiate

This scale includes items related to beginning a task or activity and independently generating ideas, responses, or problem-solving strategies. Difficulty initiating is often reported in children with brain injuries, children treated for cancer or other medical conditions (e.g., Berg et al., 2012), developmental disorders such as ASD (e.g., Kenworthy et al., 2014), and other disorders, particularly impaired cognitive speed (Langberg et al., 2014).

Development of the Dutch BRIEF-2

The Dutch BRIEF-2 (Huizinga & Smidts, 2020) is an adaptation of the American original. After approval of the publisher, we translated the American BRIEF-2. When translating items for the Dutch version, we aimed to formulate items positively, without double denials. Positive wording makes it easier for the informant to fill out the questionnaire, as it is immediately clear what is meant by the behavior. In addition, existing double denials were eliminated, which did occur in the earlier BRIEF because of literal translations from English to Dutch, leaving room for frequent questions and comments by users. From a treatment perspective, positively formulated items refer to the strengths of a child or adolescent (i.e., what a child or adolescent is able to do). This approach is in line with recent developments in the field of positive psychology and the growth mindset (e.g., Eisenberg, 2020).

Aim of the Study

The main goal of the current paper was to investigate the psychometric properties of the Dutch version of the BRIEF-2 in a representative Dutch-speaking norm sample. Using methods from classical test theory and network theory, we examined the reliability and validity of the BRIEF-2.

Method

Participants

The current samples include all participants of the normative study (parents, teachers, and adolescents), as described in the manual of the Dutch BRIEF-2 (Huizinga & Smidts, 2020). In addition to age (5–18 years), the participant selection considered sex, region, level of education, and migration background. To map the developmental trajectory of executive function and compare a child or adolescent to his/her age-related peers in the norm group, the children and adolescents were divided into seven age groups (5–6; 7–8; 9–10; 11–12; 13–14; 15–16; 17–18 years of age).

Table 1. Distribution of boys and girls within age groups, parent sample

	Boys		Girls		Total	
Age group (years)	*N*	%	*N*	%	*N*	%
5–6	106	7.1	99	6.7	205	13.8
7–8	102	6.9	111	7.5	213	14.3
9–10	97	6.5	103	6.9	200	13.5
11–12	110	7.4	99	6.7	209	14.1
13–14	110	7.4	124	8.4	233	15.7
15–16	119	8.0	107	7.2	224	15.2
17–18	97	6.5	104	7.0	201	13.5
Total	741	49.8	747	50.2	1,488	100.0

Parent Sample

The parent questionnaire was completed by 1,488 parents. Table 1 shows the distribution of the participants by age and gender.

Table 2 shows the distribution of participants in the parent sample by education level, region, and migration background relative to the national distribution according to the Statistics Netherlands (CBS, 2020). The distribution of primary and secondary education within the norm sample corresponds well to the national percentages. Relatively fewer students are included in the norm group who attend secondary vocational education, higher professional education, or university education.[2] The reason for this is the fact that in the current sample, only the first years participate (the questionnaire is for youngsters up to the age of 18), whereas the CBS numbers apply to the entire student population. The distribution across regions in the Netherlands is reasonably close to the national distribution in the norm sample, with the regions in the norm sample being distributed somewhat more evenly compared to the CBS figures. The percentage of children or adolescents with a migration background (i.e., including individuals born abroad, as well as those born in the Netherlands and at least

[2] In the Netherlands, the education system is divided in primary education (4–12 years of age) and secondary education (12–18 years of age). Secondary education is divided in several levels, corresponding to the (cognitive) abilities of youth: prevocational secondary education (vmbo), senior general secondary education (havo), preuniversity education (vwo), secondary vocational education (mbo), higher professional education (hbo), and university education (wo).

Table 2. Demographic characteristics of participants in the parent sample

Characteristics	Norm sample	The Netherlands, 2019/2020 (CBS, 2020)
Education	42.3%	38.6%
Primary education		
Secondary education	34.3%	26.2%
Prevocational secondary education	*13.3%*	
Senior general secondary education	*9.9%*	
Preuniversity education	*11.0%*	
Secondary vocational education	9.9%	13.9%
Higher professional education and university education	4.4%	21.3%
Unknown	9.1%	0
Region		
North	21.5%	9.9%
East	21.8%	21.1%
South	21.9%	21.1%
West	29.9%	47.8%
Unknown	4.9%	0
Migration background		
Yes	3.6%	24.4%
No	66.6%	75.6%
Unknown	29.8%	0

Note. Secondary education in the Netherlands is subdivided in prevocational, senior general, and preuniversity education. The numbers in these categories (in italics) refer to the percentage of the entire sample.

one of whose parents was an immigrant) is lower within the current sample compared to the national figures. This can possibly be explained by the fact that a parent with a migration background has a lower command of the Dutch language and are (likely because of the language difficulty) more difficult to reach, which makes it more difficult to fill out a questionnaire. In addition, the origin of about 30% of the participants is unknown, or the parents were unwilling to reveal. It is possible that the percentage of participants with a migration background would be higher if these data were known.

Teacher Sample

The teacher questionnaire was completed by 1,611 teachers. Table 3 shows the distribution of the participants by age and gender.

Table 4 shows the distribution of children in the teacher sample by education level, region, and migration background relative to the national distribution according to Statistics Netherlands (CBS, 2020). The distribution across the country is shown in vacation regions (North, Central, South).[3] The distribution over the different regions is almost the same. Unfortunately, due to an error in the data collection, it proved impossible to show the different levels of secondary education. The proportion of youth with a migration background was lower than the national figures. However, for about 17% of the participants, it was not known whether he or she has a migration background, or the teacher did not reveal this information.

Table 3. Distribution of boys and girls within age groups, teacher sample

	Boys		Girls		Total	
Age group (years)	*N*	%	*N*	%	*N*	%
5–6	138	8.6	133	8.3	271	16.8
7–8	109	6.8	99	6.1	208	12.9
9–10	134	8.3	108	6.7	242	15.0
11–12	137	8.5	119	7.4	256	15.9
13–14	123	7.6	120	7.4	243	15.1
15–16	129	8.0	119	7.4	248	15.4
17–18	70	4.3	73	4.5	143	8.9
Total	840	52.1	771	47.9	1,611	100.0

[3] To help authorities better manage the holiday traffic in the Netherlands, holidays are staggered across three regions: North, Central, and South. For the parent and self-report sample, we used a different classification dividing the Netherlands in four regions (following Statistics Netherlands). It was unfortunately not possible to recode the teacher sample into the four-region sample as the provinces were not recorded.

Table 4. Demographic characteristics of participants in the teacher sample

Characteristics	Norm sample	The Netherlands 2019/2020 (CBS, 2020)
Education	Unknown	38.6%
Primary education		
Secondary education	Unknown	26.2%
School holiday region		
North	34.4%	Unknown
Central	36.4%	Unknown
South	29.0%	Unknown
Unknown	0.2%	
Migration background		
Yes	8.4%	24.4%
No	74.9%	75.6%
Unknown	16.6%	0

It is possible that the percentage of participants with a migrant background would be higher if these data were available.

Self-Report Sample

The self-report questionnaire was completed by 542 adolescents. Table 5 shows the distribution of the participants by age and gender.

Table 6 shows the distribution of children in the self-report sample by education level, region, and migration background relative to the national distribution according to Statistics Netherlands (CBS, 2020). The percentage of children in primary education is smaller because in the norm sample, only children from the last year of primary education participate, whereas the CBS figures concern all children in primary education. The percentage of participants in secondary education is lower than the national figures. The CBS figures include all students; most students fall, however, outside the age limits of this questionnaire. The distribution across regions in the Netherlands approaches the national distribution very closely. The data on migration background correspond very well with the CBS figures in terms of no migration background. The proportion of children or young people with a migration background was lower than the national figures. However, for about 20% of the participants, it was not known whether he or she had a migration background or did not want to reveal. Possibly, the percentage of participants with a migration background would be higher if these data were known.

Table 5. Distribution of boys and girls within age groups, self-report sample

	Boys		Girls		Total	
Age group (years)	*N*	%	*N*	%	*N*	%
11–12	50	9.2	57	10.5	107	19.7
13–14	87	16.1	104	19.2	191	35.2
15–16	81	14.9	83	15.3	164	30.3
17–18	37	6.8	43	7.9	80	14.8
Total	255	47.0	287	53.0	542	100.0

Table 6. Demographic characteristics of participants, self-report sample

Characteristics	Norm sample	The Netherlands 2019/2020 (CBS, 2020)
Education		
Primary education	10.0%	38.6%
Secondary education	75.5%	26.2%
Prevocational secondary education	*27.1%*	
Senior general secondary education	*28.4%*	
Preuniversity education	*19.9%*	
Secondary vocational education	11.6%	13.9%
Higher professional education and university education	2.6%	21.3%
Unknown	0.4%	0
Region		
North	12.2%	9.9%
East	18.1%	21.1%
South	22.7%	21.1%
West	47.0%	47.8%
Migration background		
Yes	2.8%	24.4%
No	77.5%	75.6%
Unknown	19.7%	0

Note. Secondary education in the Netherlands is subdivided in prevocational, senior general, and preuniversity education. The numbers in these categories (in italics) refer to the percentage of the entire sample.

Procedure

Data collection took place in the two ways: through schools and through external agencies. First, through deans, school boards were approached with the request to cooperate with this study by helping to reach students (self-report), parents, or teachers. For the recruitment of the parent sample, a written description of the study was sent by e-mail to parents after approval of a school board. This e-mail explained the procedure, and after giving consent, one of the parents filled out the questionnaire online. The teacher sample too was selected after approval of the school board through an invitation by e-mail. In secondary

education, this concerned the mentors (i.e., a teacher who has a class under his/her wing and guides a student when needed; they know the students well). They completed the online questionnaire about one student: the one with number 11 on the list of their mentor class. Finally, for the adolescent sample (11–18 years old), data for the self-report were collected at the schools. After approval of the school board, the parents of the students were contacted via e-mail. This e-mail explained the procedure, and parents gave their child consent. After giving consent themselves, the students then completed the questionnaire at school during a mentor hour on study skills in a computer laboratory. The data collection took place pseudoanonymously; data were transferred encrypted.

Second, data were collected through two external agencies specializing in population data collection. With the help of these agencies, the data were supplemented to create sufficiently large samples. For the parent sample, parents of children in the appropriate age categories were selected and approached by e-mail. Self-report data were collected in the same way, by asking parents if their children (11–18 years old) would be willing to fill out the questionnaire. To collect data for the teacher sample, teachers (mentors) in both primary and secondary schools were contacted by e-mail and asked to complete the questionnaire about the 11th student on their mentor class list. Again, representativeness of the sample was always considered. Data collection via the agencies also took place pseudoanonymously.

Materials

The items of the BRIEF-2 pertain to specific everyday behavior, relevant to executive function. Parents and teachers were asked to indicate how often their child or student displayed a given behavior in the past six months by endorsing one of three responses (never, sometimes, or often). The self-report asked adolescents themselves to indicate how often they experienced a given behavior in the past six months. Response categories were also never, sometimes, or often.

BRIEF-2 Clinical Scales

Table 7 includes a brief description of the clinical scales and the number of items per clinical scale, for the parent and teacher questionnaire and the self-report. The different item numbers for each version are also evident in the original BRIEF-2 (parent and teacher questionnaires 63 items; self-report 55 items). The authors selected the items with the best predictive value to a scale. We adopted the same approach and selected the items with a value of $\alpha > .70$. As mentioned in the Introduction, we paid specific attention to increased similarity between the three versions to make them easier to compare. This process resulted in slight differences in the number of items per scale. The items of the BRIEF-2 are categorized into eight clinical scales: Inhibit, Shift, Working Memory, Self-Monitor, Task Behavior, Emotion Regulation, Plan/Organize, and Initiate. All sum scores on the clinical scales are transformed into *T*-scores.

BRIEF-2 Index and Total Scores

Based on the clinical scales, three combined indices and a total score can be calculated.[4] The content and interpretation of the three indices and the total score are discussed in the following sections.[5] All sum scores on the composite scores are transformed into *T*-scores.

Behavioral Regulation Index (BRI)

The BRI represents the ability to effectively regulate and monitor behavior. It is composed of the Inhibition and Self-Monitor scales. Behavioral regulation will generally precede cognitive regulation, thereby ensuring that cognitive regulation processes can support active problem-solving and more generally appropriate self-regulation (Gioia et al., 2015).

Emotion Regulation Index (ERI)

The ERI represents the ability to regulate emotional responses, including in response to changing situations. The ERI is composed of the Shift and Emotion Regulation scales. Appropriate emotion regulation precedes effective cognitive regulation (Perlman et al., 2014; Zelazo & Cunningham, 2007).

Cognitive Regulation Index (CRI)

The CRI represents a child's ability to control and manage cognitive processes and solve problems effectively. The CRI is composed of the following scales: Initiate, Working Memory, Plan/Organize, and Task Behavior. Appropriate cognitive regulation is required for solving complex

[4] Using exploratory and confirmatory factor analyses, Gioia et al. (2015) obtained a three-factor solution, including the Behavior Regulation Index (BRI), the Emotion Regulation Index (ERI), and the Cognitive Regulation Index (CRI). To stay as close as possible to the original BRIEF-2 – and to its translations that appeared in e.g., Denmark, Italy, and Sweden – we included these indexes in the Dutch BRIEF-2.

[5] The BRIEF-2 also contains three validity scales (Negativity, Inconsistency, and Infrequency) that provide insight into how the questionnaire was completed. As these scales do not reflect clinical relevance, we did not perform psychometric analyses on these scales.

Table 7. Description and the number of items per clinical scale, for the parent and teacher questionnaire and the self-report

Clinical scales	Parent	Teacher	Self-report	Brief description and example questions
Inhibit	9	8	7	Inhibition of behavior, thoughts, and emotions Parents & Teachers: Talks at wrong times Self-report: I talk at wrong times
Shift	8	9	8	Adapting to a change Parents & Teachers: Can find different ways to solve a problem Self-report: I can find different ways to solve a problem
Working memory	8	8	7	Processing information and holding it for a short time Parents & Teachers: Needs repetition of explanation, otherwise it does not stick Self-report: I need repetition of explanation, otherwise it does not stick
Self-monitor	6	6	4	Noticing influence of own behavior on others Parents & Teachers: Knows how his/her behavior comes across in a group Self-report: I know how my behavior comes across in a group
Task behavior	8	8	6	Keeping track (monitoring) of behavior on tasks Parents & Teachers: Checks own work for errors Self-report: I check my own work for errors
Emotion regulation	7	7	7	Modulating emotional responses Parents & Teachers: Changes mood quickly Self-report: My mood changes quickly
Plan/Organize	7	7	8	Thinking ahead and organizing information Parents & Teachers: Thinks ahead Self-report: I think ahead
Initiate	6	5	5	Getting started on a task or activity Parents & Teachers: Finds it difficult to start schoolwork on his own Self-report: I find it difficult to start schoolwork on his own

problems, learning and retrieving complex information, and being able to apply knowledge strategically (Gioia et al., 2015).

Total Score

The total score is an overall score that includes all clinical scales of the BRIEF-2. Although an assessment of the individual scale scores and the associated profile is likely to provide the most useful information, the total score may be useful as a summary measure.

Data Analysis

We investigated the reliability and validity of the BRIEF-2 in several ways, which we will describe below. To provide age-specific and sex-specific outcomes of the BRIEF, we distinguish between seven age groups (5–6; 7–8; 9–10; 11–12; 13–14, 15–16; 17–18 years of age) and boys and girls.

Reliability

The reliability of the Dutch BRIEF-2 was examined by investigation of its internal consistency, test–retest reliability, and inter-rater reliability.

Internal Consistency

We used Cronbach's α as an indication for internal consistency. Cronbach's α is based on the average correlation between items and depends on the number of items over which α is calculated (Cronbach, 1951). The internal consistency of the eight scales, three indices, and the total score of the BRIEF-2 was calculated within the three norm samples. A subgroup of items is considered a reliable measure of the construct being measured if the value of Cronbach's α is greater than .70; a value of .80 is considered good; a value of .90 or higher is considered highly reliable (Nunally & Bernstein, 1994).

Test–Retest Reliability

To determine the degree of score consistency, the three questionnaires of the BRIEF-2 were administered twice, with the second administration four to six weeks after the first administration. Test–retest reliability was examined within a subset of the normative samples of the parent questionnaire (N = 46), teacher questionnaire (N = 225), and self-report (N = 61). Test–retest reliability was determined by means of the intraclass correlation coefficient (ICC), which represents the correlation between the observations at the two measurement points. An ICC of less than 0.2 is considered very low; from 0.2 to 0.4 as

low; from 0.4 to 0.6 as moderate; from 0.6 to 0.8 as high; from 0.8 to 1.0 as very high (Landis & Koch, 1977). In addition, a paired samples *t*-test was done to compare the averages of the two measurement occasions with each other. Stability of mean scores over time is important for clinical purposes, such as monitoring recovery or response to treatment.

Inter-Rater Reliability

The inter-rater reliability of the BRIEF-2 was examined with parents and their children ($N = 98$). It is important to look at the similarities and differences between the BRIEF-2 versions because it may provide information about the different contexts from which they are completed. The inter-rater reliability between, for example, parent and teacher questionnaires of the same child is generally lower (i.e., .30–.50) than parent–parent questionnaires and teacher–student questionnaires because the home and school situations may differ in, for example, the structure offered, workload, expectations regarding behavior, and social pressures (Achenbach et al., 1987; De Los Reyes & Kazdin, 2005). The inter-rater reliability was indicated by Pearson correlation (*r*). A commonly accepted rule of thumb is that a correlation is low if the value of *r* varies around .1, medium if *r* varies around .30, and large if *r* is > .50. (Cronbach, 2004). Effect sizes are expressed using Cohen's *d*; a commonly used interpretation is to refer to effect sizes as small ($d = 0.2$), medium ($d = 0.5$), and large ($d = 0.8$; Cohen, 1988).

Validity

The Dutch BRIEF-2 is based on models widely accepted in the literature that interpret executive function as distinct but not independent processes (Gioia et al., 2000; Karr et al., 2018). When investigating its construct validity, we used the eight-factor structure of the original BRIEF-2. We examined the correlations between the scales and performed psychometric network analyses in the three norm samples.

Correlations Between Scales

Pearson correlations between scales were examined. We hypothesized that scales of theoretically related constructs (e.g., Working Memory and Plan/Organize) are more strongly correlated with each other than scales for less-related constructs (e.g., Emotion Regulation and Initiate).

Network Analyses[6]

Psychometric network analysis (e.g., Cramer et al., 2012) was used examine the network structure of the items. Recently, network analysis has made its appearance in (clinical) psychology and individual differences in cognitive performance. Here, one conceptualizes the correlational structure among the measured traits. Hence, for example, the co-occurrence of depressive symptoms such as fatigue and difficulty concentrating are seen as the consequence of direct interactions between those traits (symptoms) and not as a reflection of an underlying latent trait (such as depression). The occurrence of a depressive episode is then seen as an emergent property of interacting symptoms. The relations among these symptoms are mapped according to network models. Thus, a network model reflects direct relations between symptoms. The variables in psychometric networks are termed nodes, and the connections (usually partial correlations; edges) are usually depicted graphically to ease the network's interpretation (Cramer et al., 2012; Epskamp et al., 2018). The networks presented below were created using the qgraph package in the program R (Epskamp et al., 2012). In qgraph, the default method to depict is as a so-called weighted network, where the thickness of the edges and relative positions of the nodes depend on the strength of the statistical relations between the components (per default partial correlations). In the figures presented here, the thickness of the lines thus varies to represent the strength of the connection between the variables. To summarize the properties of the network structure, we used the recommended index strength. This indicates (numerically) how strong the connections are of a given node within the network are. A variable (node) with high strength is one that affects many variables or is affected by many variables (or both). In addition to the point estimates of the index strength of a node, we report its corresponding bootstrapped mean (and *SD*) and the nonparametric 2.5% and 97.5% quantiles. As the sample sizes would become too small when performing separate network analyses per age group and sex, we pooled the data in two groups (boys and girls) per questionnaire.

Analysis of Variance

Age group differences and sex differences were investigated by a multivariate variance analysis (MANOVA). The dependent variables included the eight clinical scales, in addition to the BRI, CRI, ERI, and Total Score age group (5–6; 7–8; 9–10; 11–12; 13–14, 15–16; 17–18 years of age), and sex (boys and girls) was included as between-subjects

[6] Measurement invariance using exploratory or confirmatory factor analysis cannot be tested with the current setup, as there is no one-to-one overlap between the instruments. Note that this essentially benefits the construct validity as this increases the sampling domain (in addition to controlling for context). Second, the number of items across questionnaires differ, leaving testing for configural measurement invariance not possible. As there are differences in length of each questionnaire, the number of latent variables to be extracted is not necessarily the same.

factors. We expressed effect sizes using partial η^2. The partial η^2 represents the proportion of the effect and error variance that is attributable to the effect (e.g., Tabachnick & Fidell, 2019). According to Cohen (1992), a value of .01, .06, and .14 represent small, medium-sized, and large effects, respectively.

Results

Reliability

Internal Consistency

Table 8 shows the mean scores, *SD*s, and internal consistency (Cronbach's α) for the BRIEF-2 scales, indices, and total score for all three versions. For an impression of the developmental course of the scores, the mean scores by age group and sex for the parent and teacher questionnaire and the self-report are shown in the Tables E1, E2, and E3 in the Electronic Supplementary Material 1 [ESM 1]. These tables also show the internal consistency for the age groups within the norm groups. The value of Cronbach's α ranged from reliable to highly reliable for the scales and indices of the parent and teacher questionnaire, and self-report, with values ranging between .60 and .97.

Test–Retest Reliability

The means and *SD*s of both measurements for all three versions are shown in Table 9. Tables E4, E5, and E6 in ESM 1 show the ICCs and paired samples *t*-test results. The ICCs on the parent and teacher questionnaires and self-reports were very high (range = .82–.96). High ICCs were observed at the Behavioral Regulation Index on the parent questionnaire (.67) and the Self-Monitor scale of the self-report (.72). The comparison of mean scores generally showed no significant differences ($p > .05$). Exceptions concerned the scales Self-Monitor of the parent questionnaire and Inhibit, Working Memory, and Emotion Regulation of the self-report, showing significant differences on the Emotion Regulation index and total score. Given the very high ICCs and very small absolute differences on these scales, this pattern might be interpreted as coincidence. However, it might also be an indication of less well-developed introspection skills in this sample. This resembles the general description of adolescence as a period including heightened sensation seeking and immature self-regulation and a decrease in self-monitoring skills (Casey, 2015; Lyons & Zelazo, 2011; Steinberg et al., 2018). Overall, based on the stable results, the BRIEF-2 can be administered repeatedly with minimal variation in scores.

Inter-Rater Reliability

Table 10 reports the correlations between the scales and indexes of the parent questionnaire and the self-report. The correlations between the scales were large and in general significant (except for the Task Behavior scale), with a mean correlation of .61. The mean effect size was −.15 (range = −.36 to .20), indicating a medium negative effect (scores on the self-report were generally slightly higher than those on the parent questionnaire).

The correlations between indexes were medium and significant, with a mean correlation of .68. The mean effect size was −0.25 (range = −.23 to −.26), indicating a small negative effect (i.e., self-report scores were higher than those on the parent questionnaire). The correlation for the total score (.72) was significant; the effect size was −.30, indicating a small negative effect.

Table 8. Means, *SD*s, and Cronbach's α for the parent and teacher questionnaire and the self-report

	Parent questionnaire N = 1,488			Teacher questionnaire N = 1,611			Self-report N = 542		
Scale/Index	*M*	*SD*	α	*M*	*SD*	α	*M*	*SD*	α
Inhibit	15.55	3.61	.80	14.39	4.36	.88	12.20	2.63	.88
Shift	13.70	3.47	.82	15.55	4.15	.85	13.27	2.85	.71
Working memory	12.66	3.38	.81	13.73	4.02	.86	11.42	3.11	.83
Self-monitor	12.93	2.77	.81	10.97	3.11	.85	7.27	2.00	.78
Task behavior	13.97	3.53	.79	13.97	4.19	.87	9.48	2.51	.74
Emotion regulation	11.46	3.21	.85	10.96	3.20	.82	11.14	2.85	.81
Plan/Organize	12.16	3.08	.76	12.59	3.51	.83	13.62	3.21	.75
Initiate	10.34	2.89	.81	9.07	3.24	.91	8.36	2.25	.73
Behavior Regulation Index	28.47	3.77	.60	25.35	6.93	.90	19.47	3.53	.71
Emotion Regulation Index	25.16	5.99	.89	26.51	6.68	.89	24.41	4.89	.82
Cognitive Regulation Index	49.12	11.48	.93	49.36	13.78	.96	42.89	9.57	.92
Total score	102.76	17.38	.93	101.23	24.26	.97	86.77	15.61	.93

Table 9. Means, *SD*s at Time 1 (test T1) and Time 2 (retest T2)

	Parent questionnaire $N = 46$		Teacher questionnaire $N = 226$		Self-report $N = 61$	
Scale/Index	*M*	*SD*	*M*	*SD*	*M*	*SD*
Inhibit – T1	14.72	3.15	14.48	4.47	12.95	2.70
Inhibit – T2	14.50	3.33	14.32	4.35	13.93	2.55
Shift – T1	13.93	4.14	15.52	4.21	13.85	2.88
Shift – T2	14.24	3.86	15.38	4.33	14.39	2.92
Working memory – T1	12.67	3.33	13.77	4.30	12.34	3.03
Working memory – T2	13.00	3.42	13.51	4.31	13.03	3.01
Self-monitor – T1	12.67	2.70	10.96	3.17	7.67	1.95
Self-monitor – T2	13.70	2.49	10.72	3.16	7.28	1.60
Task behavior – T1	13.26	4.05	14.08	4.18	10.51	2.49
Task behavior – T2	13.48	4.03	13.92	4.27	10.51	2.55
Emotion Regulation – T1	11.26	3.09	11.06	3.29	11.93	2.99
Emotion Regulation – T2	11.15	3.26	10.97	3.27	12.80	2.83
Plan/Organize – T1	12.70	3.55	12.64	3.77	15.30	3.09
Plan/Organize – T2	12.07	3.51	12.44	3.92	15.34	3.24
Initiate – T1	10.52	3.03	9.04	3.22	9.30	2.36
Initiate – T2	10.22	2.95	8.89	3.32	9.49	2.38
Behavior Regulation Index – T1	27.39	2.84	25.54	7.20	20.62	3.28
Behavior Regulation Index – T2	28.20	2.73	25.29	7.08	21.21	3.15
Emotion Regulation Index – T1	25.20	6.61	26.58	6.83	25.79	4.97
Emotion Regulation Index – T2	25.39	6.43	26.35	6.98	27.20	5.10
Cognitive Regulation Index – T1	49.15	13.09	49.53	14.46	47.44	9.73
Cognitive Regulation Index – T2	48.76	12.80	48.76	14.79	48.38	10.08
Total score – T1	101.74	19.39	101.54	25.74	93.85	15.99
Total score – T2	102.35	18.87	100.15	26.09	96.79	16.03

Note. – T1 = measurement 1; – T2 = measurement 2.

Table 10. Inter-rater reliability of the BRIEF-2 ($N = 98$)

		Parent questionnaire		Self-report			
Scale/Index	*r*	*M*	*SD*	*M*	*SD*	Mean difference	Cohen's *d*
Inhibit	.76*	1.75	0.40	1.84	0.42	−0.08	−0.29
Shift	.46*	1.97	0.33	2.09	0.30	−0.12	−0.36
Working memory	.73*	1.84	0.38	1.77	0.46	0.06	0.20
Self-monitor	.74*	2.14	0.38	2.18	0.50	−0.04	−0.13
Task behavior	.17	2.03	0.27	2.16	0.30	−0.13	−0.36
Emotion Regulation	.79*	1.72	0.46	1.71	0.44	0.01	0.02
Plan/Organize	.55*	2.03	0.28	2.12	0.31	−0.09	−0.33
Initiate	.71*	1.93	0.40	1.97	0.44	−0.04	−0.13
Behavior Regulation Index	.69*	1.91	0.24	1.96	0.30	−0.05	−0.25
Emotion Regulation Index	.68*	1.86	0.35	1.92	0.30	−0.06	−0.23
Cognitive Regulation Index	.70*	1.95	0.27	2.01	0.25	−0.05	−0.26
Total score	.72*	1.92	0.24	1.97	0.24	−0.05	−0.30

Note. *significant at $p = .001$; two-tailed.

Validity

Correlations Between the Scales

Tables 11, 12, and 13 show the correlations between the scales for the norm sample of the parent and teacher questionnaire and the self-report, broken down for boys and girls. The scales on the BRIEF-2 are moderately to strongly and significantly correlated. On the parent questionnaire, the range is .22–.90, noting that the Self-Monitor scale correlates negatively with the other scales (range = −.22 to .39). The range of the teacher questionnaire and the self-report ranged from .42 to .80 and .18 to .75, respectively. On the self-report, in the girl sample, three correlations were not significant (Monitoring with Working Memory, Inhibition, and Emotion Regulation).

Network Analyses

The upper left and right panels of Figure 1 and Table E7 in ESM 1 show the results of the psychometric network analysis on the items from the parent questionnaire and separately for boys and girls. The networks show that the edges range from strong to very strong. The three items with the highest strength are starts a task or chore and gets stuck, due to lack of overview; needs prompting to start a task, even if the will is there; feels how his/her behavior comes across to others in the group of boys and thinks ahead; stays stuck in the same thought; and has difficulty with transitions (activities, situations, etc.) in the group of girls. This variation results in clustering of items. Clustering is in such a way that the items map well according to the intended scales. For example, items that indicate the higher-order factor (scale) Plan/Organize are close together. Yet, items within a given cluster also tend to have several connections with items outside that cluster, indicating the typically observed interdependence of executive functions.

The middle left and right panels of Figure 1 and Table E8 in ESM 1 show the results of the psychometric network analysis on the items from the Teacher questionnaire, again separately for boys and girls. The main findings are similar to those above: The edges range from strong to very strong, the items cluster according to the intended scales, and clusters are interconnected. The three items with the highest strength for the boys are as follows: finds it difficult to start schoolwork on his/her own; forgets what he/she was doing; and forgets where he/she left off. For girls, they are as follows: finds new situations very exciting; does without thinking; and thinks ahead. Remarkable in this analysis is that for both boys and girls, the item cries quickly has no edges with the other items.

The lower left and right panel of Figure 1 and Table E9 in ESM 1 show the results of the psychometric network analysis on the items from the self-report, again separately for boys and girls. The main findings are to a certain extent similar to those above, but noticeable differences also exist: The edges range from strong to very strong, some of

Table 11. Correlations of the scales of the BRIEF-2, parent questionnaire (N = 1,488)

Scale/Index	Inhibit	Shift	Working memory	Self-monitor	Task behavior	Emotion regulation	Plan/Organize	Initiate	Behavior Regulation Index	Emotion Regulation Index	Cognitive Regulation Index
Inhibit	—	.50	.56	−.38	.51	.60	.52	.47	.71	.61	.58
Shift	.40	—	.56	−.39	.41	.62	.61	.58	.20	.91	.60
Working memory	.52	.47	—	−.29	.69	.51	.77	.71	.33	.60	.90
Self-monitor	−.25	−.26	−.27	—	−.31	−.32	−.40	−.31	.38	−.39	−.37
Task behavior	.50	.41	.72	−.30	—	.42	.73	.67	.28	.46	.88
Emotion Regulation	.50	.60	.41	−.22	.36	—	.51	.48	.36	.89	.54
Plan/Organize	.43	.56	.74	−.33	.73	.45	—	.77	.22	.62	.91
Initiate	.40	.48	.70	−.25	.66	.40	.73	—	.23	.59	.88
Behavior Regulation Index	.73	.18	.28	.48	.24	.30	.15	.19	—	.30	.30
Emotion Regulation Index	.50	.90	.49	−.27	.43	.89	.56	.49	.27	—	.64
Cognitive Regulation Index	.53	.54	.89	−.33	.89	.45	.90	.86	.25	.56	—

Note. Correlations for boys (N = 741) above the diagonal, correlations for girls (N = 747) below the diagonal. All correlations are significant with $p < .001$.

Table 12. Correlations of the scales of the BRIEF-2, teacher questionnaire (*N* = 1,693)

Scale/Index	Inhibit	Shift	Working memory	Self-monitor	Task behavior	Emotion regulation	Plan/ Organize	Initiate	Behavior Regulation Index	Emotion Regulation Index	Cognitive Regulation Index
Inhibit	—	.50	.52	.48	.61	.67	.51	.50	.94	.63	.59
Shift	.50	—	.63	.54	.59	.65	.70	.60	.61	.93	.69
Working memory	.61	.58	—	.45	.79	.40	.83	.78	.51	.58	.93
Self-monitor	.49	.53	.48	—	.51	.46	.54	.42	.52	.56	.52
Task behavior	.69	.55	.80	.55	—	.48	.77	.75	.61	.59	.91
Emotion Regulation	.65	.64	.44	.43	.49	—	.46	.39	.89	.88	.47
Plan/Organize	.57	.72	.82	.56	.78	.50	—	.81	.54	.66	.93
Initiate	.59	.64	.78	.47	.76	.48	.81	—	.49	.56	.90
Behavior Regulation Index	.93	.61	.59	.51	.66	.88	.59	.59	—	.80	.59
Emotion Regulation Index	.62	.93	.58	.54	.58	.87	.69	.63	.81	—	.65
Cognitive Regulation Index	.67	.67	.93	.56	.92	.52	.93	.90	.67	.67	—

Note. Correlations for boys (*N* = 881) above the diagonal, correlations for girls (*N* = 812) below the diagonal. All correlations are significant with $p < .001$.

Table 13. Correlations of the scales of the BRIEF-2, self-report (*N* = 542)

Scale/Index	Inhibit	Shift	Working memory	Self-monitor	Task behavior	Emotion regulation	Plan/ Organize	Initiate	Behavior Regulation Index	Emotion Regulation Index	Cognitive Regulation Index
Inhibit	—	.39	.56	.18	.46	.52	.54	.49	.84	.52	.58
Shift	.26	—	.55	.42	.51	.54	.62	.53	.52	.88	.63
Working memory	.57	.35	—	.29	.66	.46	.75	.68	.57	.58	.89
Self-monitor	*.10*	.39	*.08*	—	.44	.18	.43	.19	.69	.34	.39
Task behavior	.46	.34	.55	.30	—	.32	.77	.64	.58	.48	.87
Emotion Regulation	.49	.40	.55	*.06*	.36	—	.41	.34	.48	.87	.43
Plan/Organize	.45	.44	.62	.30	.61	.45	—	.74	.64	.59	.93
Initiate	.48	.43	.65	.19	.49	.42	.68	—	.46	.50	.85
Behavior Regulation Index	.82	.42	.48	.65	.52	.41	.51	.47	—	.57	.64
Emotion Regulation Index	.45	.83	.54	.27	.42	.84	.53	.51	.50	—	.61
Cognitive Regulation Index	.59	.46	.85	.25	.77	.54	.88	.83	.59	.60	—

Note. Correlations for boys (*N* = 255) above the diagonal, correlations for girls (*N* = 287) below the diagonal. All correlations are significant with $p < .001$, except when in italics.

the items cluster according to the intended scales, and clusters are interconnected, but some items do not cluster that well. For example, items that are intended to indicate Initiate and Task Behavior are somewhat scattered. In addition, for the boys, the items "I immediately do what comes into my mind," "I tend to focus on details (and lose

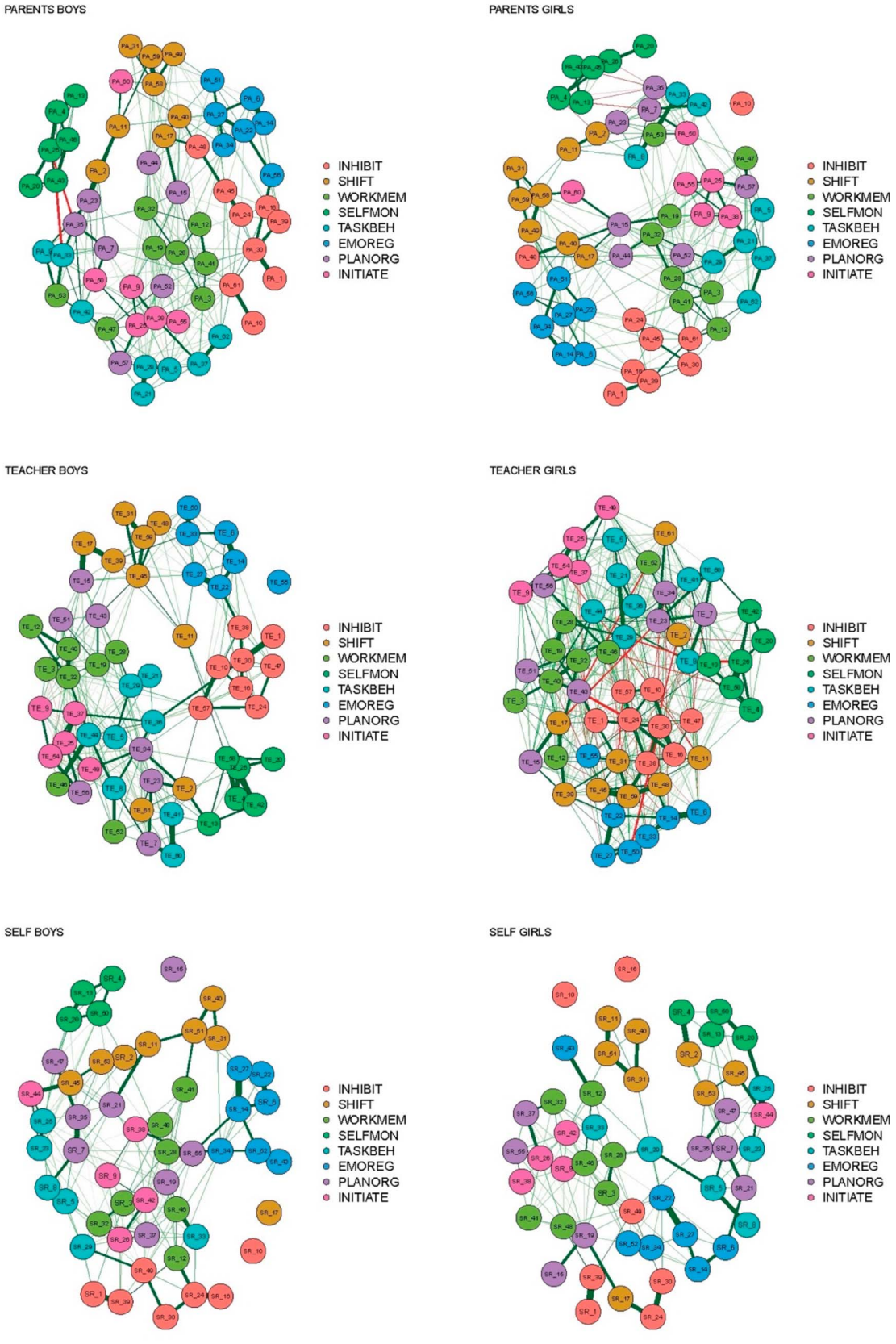

Figure 1. Network structures of the BRIEF-2 parent, teacher, and self-report, for boys and girls apart. WorkMem = Working Memory; SelfMon = Self Monitor; TaskBeh = Task Behavior; EmoReg = Emotion Regulation; PlanOrg = Plan/Organize.

the overview);" and "I get stuck on a subject or activity have no edges with the other items." For the girls, this applies to the items "I immediately do what comes into my mind" and "I act wilder than my classmates." The three items with the highest strength for the boys are "I think ahead;" "I forget where I left off;" and "I do things without thinking carefully." For the girls they are "I want to start a task, but I can't;" "I start a task or chore and get stuck because of the lack of overview;" and "I do wilder than my classmates."

Analysis of Variance

Tables E1, E2, and E3 in ESM 1 show the mean scores and *SD*s with respect to age and sex at the parent and teacher questionnaires and the self-report, respectively.

Parent Questionnaire

The results of the MANOVA showed that there are significant differences in means between age groups, Wilks $\Lambda = .80$, $F(54, 7{,}479.76) = 6.19$, $p < .001$; partial $\eta^2 = .04$. Significant effects were found at all scales and indices (p's < .001; Shift $p = .002$). Post hoc testing (using a Bonferroni correction) showed the following results at the different scales:

Inhibition scale: The 5–6-year-olds had a significantly higher score than the subsequent age groups; the 7–8-year-olds had significantly higher scores on the 13+ age groups; the 9–13-year-olds had higher scores than the 15+ age groups. **Shift scale**: The 5–8-year-olds and 11–12-year-olds had significantly higher scores than the 17–18-year-olds. **Working Memory scale**: The 5–14-year-olds had significantly higher scores than the 17–18-year-olds. **Monitor scale**: the 5–6-year-olds had significantly lower scores than the 9–10-year-olds and 13–18-year-olds; the 7–8-year-olds had lower scores than the 13–14-year-olds. **Task Behavior scale**: The 17–18-year-olds had significant lower scores than the 5–16-year-olds. **Emotion Regulation scale**: The 5–6-year-olds had significantly higher scores than the 13+ age groups; the 7–8-year-olds had higher scores than the 15+ age groups; the 9–10-year-olds had higher scores than the 17–18-year-olds. **Plan/Organize scale**: The 11–14-year-olds had significantly higher scores than the 17-18-year-olds. At the Initiate, the 9–16-year-olds had significantly higher scores than the 17–18-year-olds. **Behavior Regulation Index**: The 17–18-year-olds had significant lower scores than the preceding age groups; the 15–16-year-olds had significant lower scores than the 5–9-year-olds. **Emotion Regulation Index**: The 5–6-year-olds had significantly higher scores than the 13+ age groups; the 7–12-year-olds had higher scores than the 17–18-year-olds. **Cognitive Regulation Index**: The 17–18-olds had significantly lower scores than the preceding age groups.

In addition, a main effect was found for sex, Wilks $\Lambda = .94$, $F(9, 1{,}466) = 9.79$, $p < .001$; partial $\eta^2 = .06$. Boys showed significantly higher scores compared to girls on all scales and indices. The interaction between age and sex was significant, Wilks $\Lambda = .95$, $F(54, 7{,}479.76) = 1.55$, $p = .006$; partial $\eta^2 = .01$. Tests of between-subject effects showed significant interactions at the Emotion Regulation scale ($p = .006$) and the Emotion Regulation Index ($p = .008$), with boys showing higher scores than girls but a general decrease when they grow older where girls showed a general increase of scores.

Teacher Questionnaire

The results of the MANOVA showed that there are significant differences in means between age groups, Wilks $\Lambda = .81$, $F(54, 8{,}106.94) = 6.29$, $p < .001$; partial $\eta^2 = .03$. Significant effects were found at the Monitor ($p < .001$), Plan/Organize ($p = .024$), and the Initiate scales ($p < .001$), as well as at the Cognitive Regulation Index ($p = .006$). Post hoc testing (using a Bonferroni correction) showed at the Monitor scales significantly higher scores of 5–6-year-olds compared to the subsequent age groups. The 7–8-year-olds showed significantly higher scores compared to 11–12-year-olds and 15–18-year-olds. At the Initiate scale, the 15–18-year-olds showed significantly higher scores compared to 5–12-year-olds. The 9–10-year-olds showed significantly lower scores than the 13–14-year-olds.

In addition, a main effect was found for sex, Wilks $\Lambda = .85$, $F(9, 1{,}589) = 32.35$, $p < .001$; partial $\eta^2 = .16$. Boys showed significantly higher scores compared to girls on all scales and indices. The interaction between age and sex was significant, Wilks $\Lambda = .95$, $F(54, 8{,}106.94) = 1.68$, $p = .001$; partial $\eta^2 = .01$. Tests of between-subjects effects showed a significant interaction of age and sex on the Emotion Regulation scale ($p = .008$) and the Emotion Regulation Index ($p = .023$), with boys showing higher scores than girls but a general decrease when they grow older where girls showed a general increase of scores.

Self-Report

The results of the MANOVA showed no significant differences in means between age groups, $p > .06$. A main effect was found for sex, Wilks $\Lambda = .91$, $F(9, 526) = 6.12$, $p < .001$; partial $\eta^2 = .10$. Boys showed significantly lower scores compared to girls on all scales and indices. The interaction between age and sex was significant, Wilks $\Lambda = .93$, $F(27, 1{,}536.83) = 1.55$, $p = .035$; partial $\eta^2 = .03$. Tests of between-subjects effects showed a significant interaction of age and sex on the Inhibition ($p = .003$), Shift ($p = .003$), Working Memory ($p = .011$), Plan/Organize ($p = .027$) scales and the Behavior ($p = .001$), Emotion ($p = .009$), and Cognitive ($p = .034$) Regulation Indices and total score ($p = .005$), with

boys showing lower scores than girls (except the 15–16-year-olds) and a general decrease of scores when they grow older.

Discussion

The BRIEF-2 (Gioia et al., 2015) is a widely used questionnaire to measure daily behavior related to executive function behaviors in the home and school environment of children between 5 and 18 years old. The current study was conducted to investigate the psychometric properties of the Dutch version of the BRIEF-2 in a representative Dutch-speaking norm sample. In addition, we examined the age-related and sex differences with respect to the factors that the BRIEF-2 aims to measure.

The findings of this study show that the internal consistency and the test–retest reliability of the parent and teacher questionnaires and the self-report of the Dutch version of the BRIEF-2 are high to very high. These results are equivalent to the original version of the BRIEF-2 (Gioia et al., 2015), indicating that the Dutch BRIEF-2 is a reliable measure of executive function. The inter-rater reliability was examined in the parent and self-report sample. The results indicated significant medium correlations and small but negative effect sizes, with higher self-report scores than parent scores. This result is equivalent to the original BRIEF-2, where the lowest correlations and effect sizes were seen between adolescent self-reports and parent questionnaires. Investigating differences between raters' observations are often clinically informative and may be meaningfully interpreted (see also Achenbach et al., 1987). Analyses indicated by classical test theory document strong internal consistency and test–retest reliabilities for the parent, teacher, and self-report forms along with appropriate inter-rater reliability for the parent and self-report forms in the norm sample. Overall, the evidence supports reliable interpretation of BRIEF-2 scores.

With regards to the construct validity, we took two different approaches, as the conceptualization of executive function during development is the topic of ongoing debate. Mixed results are reported among child/adolescent and adult samples, in general showing greater unidimensionality of executive function among child/adolescent samples and both unity and diversity among adult samples (Baggetta & Alexander, 2016; Karr et al., 2018). First, the scales on the three versions correlated moderately to strongly. Second, to gain insight in the clustering of the eight factors, we applied psychometric network analyses to the parent, teacher, and self-report samples for boys and girls separately. The results of the network analyses consistently showed that the edges (i.e., connections between individual items) range from strong to very strong, resulting in clustering of items that map well according to the higher-order factor solution. Remarkably, items within a given cluster also tend to have several connections with items outside that cluster. This finding corresponds well with the notion that executive function behavior often cannot be seen as representative for one exclusive factor (Karr et al., 2018). For example, switching flexibly to another activity often implies the stopping of the current activity (inhibition) and – depending on the situation – working memory, to update the information on the current activity. Obviously, an adult sample is not included in this study. Future research would benefit from extending the sample with adults, as this renders the unique possibility of examining the factor structure from childhood to adulthood. One complicating factor concerns the fact that the content of the BRIEF-2 items refers to situations in the child's life and for the largest part do not represent the adult life. Another important focus for future research involves establishing more insight in the organization of executive function by focusing on the network structure of individual items. Overall, we found that the factorial structure of the BRIEF-2 was consistent with our expectations. We consider these results to be an important prerequisite to further research into the construct validity of the BRIEF-2 and in a broader perspective, research into the organization of executive function during development.

In addition to examining the psychometric properties of the BRIEF-2, we examined age and sex differences. In our MANOVAs, we observed main effects of age and sex at the parent and teacher questionnaires, a main effect of sex at the self-report, as well as interactions of age and sex at all three questionnaires. Overall, at the parent and teacher questionnaire, younger children showed more behavior problems related to executive function than older children and adolescents. This is in line with a vast of literature showing an improvement in executive function over the course of childhood and adolescence (e.g., Diamond, 2013; Huizinga et al., 2006; Lehto et al., 2003; Luciana & Nelson, 2002; Steinberg et al., 2018; Zelazo & Carlson, 2012). Observations of monitoring skills showed the opposite result, with young children showing fewer problems compared with adolescents. The observations of parents and teachers at the BRIEF-2 likely reflect the crucial role of self-monitoring in the development of self-regulated behavior. Adolescence is period-characterized by heightened sensation-seeking and immature self-regulation, with adolescents showing a decrease in self-monitoring skills (Casey, 2015; Lyons & Zelazo, 2011; Steinberg et al., 2018). In addition, we found small but significant changes between boys and girls, in favor of girls. This in line with the patterns found in studies applying direct assessments of executive function (e.g., Isquith et al., 2004; Wiebe et al., 2008; Zelazo et al., 2013). An interaction effect of age and sex on the Emotion Regulation scale and index,

respectively, showed a decrease in scores when children grow older and in girls compared to boys. Although statistical comparisons regarding age and sex are lacking in the original version of the BRIEF-2 (Gioia et al., 2015), the results of the present study are at face value comparable to the results of Gioia et al., who showed a similar pattern of results.

At the self-report, we found an increased level of problems with executive function in girls compared to boys, on all scales and indices. Although significant, the differences between the two groups were very small and impossible to interpret in terms of differences in daily life behavior. This seems to relate to the general finding of no convincing evidence for sex differences in executive function (Grissom & Reyes, 2019). In their meta-analytic review, Grissom & Reyes showed that there might be individual factors that may show a tendency toward a sex bias (e.g., increased impulsive action in males, improved working memory in females). There is, however, no convincing evidence for a general sex difference in executive function. They hypothesize that girls and boys may use different mechanisms use to solve the same cognitive problems. Hence, the apparent difference between sexes is not explained by a difference in ability, but by a difference in strategy. As for the BRIEF-2, an interaction of age and sex further qualified the observed sex difference by showing a decrease of executive function problems with age. Despite the lack of statistical comparisons regarding age and sex in the original Self-Report version of the BRIEF-2 (Gioia et al., 2015), the results of the present study are at face value comparable to the results of Gioia et al.

Limitations

The main limitation of the present study concerns the separate teacher sample. This hampered the examination of BRIEF-2 scores from a single child from three perspectives. Obviously, multiple assessments of a child or adolescent's behavior provide a more comprehensive picture that can be used to identify a child's needs. Both similarities and differences between assessments provide clinically useful information. Given the large sample sizes, it turned out to be very complicated to obtain consent from the parents, teacher, and the child or adolescent at once. Nevertheless, the inclusion of teachers of 11-to-18-year-olds in the BRIEF-2 is a valuable improvement compared to the original Dutch BRIEF, for which these data were not available. Another limitation concerns the lack of data on, for example, learning and developmental disorders, traumatic brain injury, or psychiatric diseases in a Dutch sample. These differences have been established in the original version of the BRIEF-2 and clearly show the usefulness of the BRIEF-2 in establishing diagnoses and treatment plans (Gioia et al., 2015). In addition, additional research regarding the divergent validity of the BRIEF-2 is needed.

In addition, we only obtained indications of construct validity. The currents study lacks the establishment of convergent and divergent validity using, for example, task-based measures of executive function. Previous research, however, consistently indicated low associations between rating measures of executive function reports and performance-based measures of executive functions (e.g., Friedman & Banich, 2019; Toplak et al., 2013). These findings are generally interpreted as two types of measures assessing different mental constructs that independently predict (individual differences in) behavior, that is, the efficiency of cognitive abilities and success in goal pursuit, respectively.

Conclusion

Taken together, the BRIEF-2 can be considered as a valid and reliable questionnaire that provides information on the role of executive function in the child's and adolescent's functioning in the home and school environment. It is based on a strong previous measure (i.e., the BRIEF; Gioia et al., 2000) with historical use in both research and clinical settings. Its scales, indexes, and composite scores mirror the umbrella construct of executive function. Test standardization was conducted using large and robust sample sizes in parents, teachers, and adolescents. The BRIEF-2 scores have strong internal consistency and test–retest stability and moderate-to-strong construct validity. This forms a solid base for use in research and clinical settings.

Electronic Supplementary Material

The electronic supplementary material is available with the online version of the article at https://doi.org/10.1027/2698-1866/a000038

ESM 1. Tables with mean raw scores and reliabilities for parent questionnaire, teacher questionnaire, and self-report. Tables with test–retest reliabilities of the BRIEF-2 for parent questionnaire, teacher questionnaire, and self-report. Tables with network structures for boys and girls.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional-problems. Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213–232. https://doi.org/10.1037/0033-2909.101.2.213

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).

Baggetta, P., & Alexander, P. A. (2016). Conceptualization and operationalization of executive function. *Mind Brain and Education*, *10*(1), 10–33. https://doi.org/10.1111/mbe.12100

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

Beer, J. S., Lombardo, M. V., & Bhanji, J. P. (2010). Roles of medial prefrontal cortex and orbitofrontal cortex in self-evaluation. *Journal of Cognitive Neuroscience*, *22*(9), 2108–2119. https://doi.org/10.1162/jocn.2009.21359

Berg, C., Edwards, D. F., & King, A. (2012). Executive function performance on the children's kitchen task assessment ith children with sickle cell disease and matched controls. *Child Neuropsychology*, *18*(5), 432–448. https://doi.org/10.1080/09297049.2011.613813

Casey, B. J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annual Review of Psychology*, *66*, 295–319. https://doi.org/10.1146/annurev-psych-010814-015156

CBS. (2020). *Statline*. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/71450ned/table?ts=1644686979513https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70072ned/table?ts=1644687160345https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37296ned/table?ts=1644687202106

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cramer, A. O. J., Kendler, K. S., & Borsboom, D. (2012). A constructionist account of emotional disorders. *Behavioral and Brain Sciences*, *35*(3), 146–147. https://doi.org/10.1017/s0140525x11001488

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/bf02310555

Cronbach, L. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, *64*(3), 391–418. https://doi.org/10.1177/0013164404266386

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, *131*(4), 483–509. https://doi.org/10.1037/0033-2909.131.4.483

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Eisenberg, N. (2020). Considering the role of positive emotion in the early emergence of prosocial behavior: Commentary on Hammond and Drummond (2019). *Developmental Psychology*, *56*(4), 843–845. https://doi.org/10.1037/dev0000880

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4), 1–18. https://doi.org/10.18637/jss.v048.i04

Friedman, N. P., & Banich, M. T. (2019). Questionnaires and task-based measures assess different aspects of self-regulation: Both are needed. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(49), 24396–24397. https://doi.org/10.1073/pnas.1915315116

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

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). *BRIEF2: Behavior Rating Inventory of Executive Function* (2nd ed.). Psychological Assessment Resources.

Gioia, G. A., Isquith, P. K., Retzlaff, P. D., & Pratt, B. M. (2001). Modeling executive functions with everyday behaviors: A unitary or fractionated system? *Brain and Cognition*, *47*(1-2), 203–207.

Grieve, A., Webne-Behrman, L., Couillou, R., & Sieben-Schneider, J. (2014). Self-report assessment of executive functioning in college students with disabilities. *Journal of Postsecondary Education and Disability*, *27*(1), 19–32. http://www.ahead-archive.org/uploads/publications/JPED/JPED27_2/JPED27_2_FULL%20DOCUMENT.pdf

Grissom, 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. https://doi.org/10.1038/s41386-018-0179-5

Guy, S. C., Isquith, P. K., & Gioia, G. A. (2004). *Behavior Rating Inventory of Executive Function – Self-report version*. Psychological Assessment Resources.

Gyurak, A., Goodkind, M. S., Kramer, J. H., Miller, B. L., & Levenson, R. W. (2012). Executive functions and the down-regulation and up-regulation of emotion. *Cognition & Emotion*, *26*(1), 103–118. https://doi.org/10.1080/02699931.2011.557291

Hovik, K. T., Egeland, J., Isquith, P. K., Gioia, G., Skogli, E. W., Andersen, P. N., & Oie, M. (2017). Distinct patterns of everyday executive function problems distinguish children with Tourette syndrome from children with ADHD or autism spectrum disorders. *Journal of Attention Disorders*, *21*(10), 811–823. https://doi.org/10.1177/1087054714550336

Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. *Infant and Child Development*, *20*(3), 251–271. https://doi.org/10.1002/icd.736

Huizinga, M., Baeyens, D., & Burack, J. A. (2018). Editorial: Executive function and education. *Frontiers in Psychology*, *9*, Article 1357. https://doi.org/10.3389/fpsyg.2018.01357

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Huizinga, M., & Smidts, D. P. (2011). Age-related changes in executive function: A normative study with the Dutch version of the Behavior Rating Inventory of Executive Function (BRIEF). *Child Neuropsychology*, *17*(1), 51–66. https://doi.org/10.1080/09297049.2010.509715

Huizinga, M., & Smidts, D. P. (2012). *BRIEF Executieve Functies Gedragsvragenlijst: Handleiding* [BRIEF Executive Functions Behavioral Questionnaire: Manual]. Hogrefe.

Huizinga, M., & Smidts, D. P. (2020). *BRIEF-2 Vragenlijst executieve functies voor 5- tot en met 18-jarigen* [BRIEF-2 Executive Function Questionnaire for 5–18-year-olds]. Hogrefe.

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

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies [Review]. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Kenworthy, L., Gutermuth Anthony, L., Naiman, D. Q., Cannon, L., Wills, M. C., Luong-Tran, - C., Adler Werner, M., Alexander, K. C., Strang, J., Bal, E., Sokoloff, J. L., & Wallace, G. L. (2014). Randomized controlled effectiveness trial of executive function

intervention for children on the autism spectrum. *Journal of Child Psychology and Psychiatry*, *55*(4), 374–383. https://doi.org/10.1111/jcpp.12161

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Langberg, J. M., Becker, S. P., & Dvorsky, M. R. (2014). The association between sluggish cognitive tempo and academic functioning in youth with attention-deficit/hyperactivity disorder (ADHD). *Journal of Abnormal Child Psychology*, *42*(1), 91–103. https://doi.org/10.1007/s10802-013-9722-3

Lehto, J. H., Juujaervi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59–80. https://doi.org/10.1348/026151003321164627

Luciana, M., & Nelson, C. A. (2002). Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: Performance in 4- to 12-year-old children. *Developmental Neuropsychology*, *22*(3), 595–624. https://doi.org/10.1207/S15326942DN2203_3

Luria, A. R. (1966). *Higher cortical function in man*. Basic Books.

Lyons, K. E., & Zelazo, P. D. (2011). Monitoring, metacognition, and executive function: Elucidating the role of self-reflection in the development of self-regulation. *Advances in Child Development and Behavior*, *40*, 379–412. https://doi.org/10.1016/B978-0-12-386491-8.00010-4

McAuley, T., Chen, S., Goos, L., Schachar, R., & Crosbie, J. (2010). Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *Journal of the International Neuropsychological Society*, *16*(3), 495–505. https://doi.org/10.1017/s1355617710000093

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Nunally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). MacGraw-Hill.

Perlman, S. B., Hein, T. C., Stepp, S. D., & Consortium, L. (2014). Emotional reactivity and its impact on neural circuitry for attention-emotion interaction in childhood and adolescence. *Developmental Cognitive Neuroscience*, *8*, 100–109. https://doi.org/10.1016/j.dcn.2013.08.005

Shallice, T. (1982). Specific impairments in planning. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, *298*(1089), 199–209. https://doi.org/10.1098/rstb.1982.0082

Smidts, D. P., & Huizinga, M. (2009). *BRIEF Executieve Functies Gedragsvragenlijst: Handleiding* [*BRIEF Executive Functions Behavioral Questionnaire: Manual*]. Hogrefe.

Steinberg, L., Icenogle, G., Shulman, E. P., Breiner, K., Chein, J., Bacchini, D., Chang, L., Chaudhary, N., Di Giunta, L., Dodge, K. A., Fanti, K. A., Lansford, J. E., Malone, P. S., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Uribe Tirado, L. M., … Takash, H. M. S. (2018). Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation. *Developmental Science*, *21*(2), Article e12532. https://doi.org/10.1111/desc.12532

Stuss, D. T. (1992). Biological and psychological development of executive functions. *Brain and Cognition*, *20*(1), 8–23.

Suchy, Y. (2009). Executive functioning: Overview, assessment, and research issues for non-neuropsychologists. *Annals of Behavioral Medicine*, *37*(2), 106–116. https://doi.org/10.1007/s12160-009-9097-4

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2009). Executive functions: Performance-based measures and the Behavior Rating Inventory of Executive Function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology*, *15*(1), 53–72. https://doi.org/10.1080/09297040802070929

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct?. *Journal of Child Psychology and Psychiatry*, *54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*(2), 575–587. https://doi.org/10.1037/0012-1649.44.2.575

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH toolbox cognition battery (cb): Measuring executive function and attention. *Monographs of the Society for Research in Child Development*, *78*(4), 16–33. https://doi.org/10.1111/mono.12032

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

Zelazo, P. D., & Cunningham, W. A. (2007). Executive function: Mechanisms underlying emotion regulation. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 135–158). Guildford Press.

Zelazo, P. D., Müller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function: Cognitive complexity and control – revised. *Monographs of the Society for Research in Child Development*, *68*(3), 93–119. https://doi.org/10.1111/j.0037-976X.2003.00266.x

History
Received July 26, 2022
Revision received February 3, 2023
Accepted March 28, 2023
Published online June 20, 2023

Section: Methodological Topics in Assessment

Authorship
Mariëtte Huizinga: manuscript preparation and statistical analyses (except psychometric network analysis) and development of BRIEF-2. Diana Smidts: development of BRIEF-2. Dieter Baeyens: manuscript preparation. Kees-Jan Kan: psychometric network analyses.

Funding
The data were collected with financial support of Hogrefe Publishers, The Netherlands.

ORCID
Mariëtte Huizinga
https://orcid.org/0000-0001-7007-8952
Diana P. Smidts
https://orcid.org/0000-0003-3471-5728
Dieter Baeyens
https://orcid.org/0000-0002-3875-687X
Kees-Jan Kan
https://orcid.org/0000-0003-0088-9906

Mariëtte Huizinga
Department of Education and Family Studies
Faculty of Behavioral and Movement Sciences
Vrije Universiteit Amsterdam
Van der Boechorststraat 7
1081 BT Amsterdam
The Netherlands
m.huizinga@vu.nl