Product Overview

Jack A. Naglieri, Ph.D. & Sam Goldstein, Ph.D.

Please Note: All data presented in this overview are subject to change.

Overview

The Comprehensive Executive Function Inventory (CEFI™) is a 100-item rating scale designed to measure behaviors associated with executive function in children and youths aged 5 through 18 years. The rating scale is completed by a parent, teacher, or the youth. When used in combination with other information, results from the CEFI help in guiding diagnostic decisions, treatment planning, and ongoing monitoring of treatment progress.

The CEFI has parent and teacher forms for children and youth aged 5 to 18 years, and a self-report form for youth aged 12 to 18 years. All three forms have 100 items and include a Full Scale as well as nine CEFI Scales: Attention, Emotion Regulation, Flexibility, Inhibitory Control, Initiation, Organization, Planning, Self-Monitoring, and Working Memory. In addition, a Consistency Index, Negative Impression and Positive Impression scales are provided.

Key Features

- **Normative samples:** the CEFI was normed on a large national scale representative of U.S. population on a number of key demographic variables within 2% of Census targets
- **Multi-rater:** parent, teacher, and self-reports provide a comprehensive evaluation of executive function
- Evaluate executive function **strengths** and **weaknesses across nine scales**
- **Age range**: 5–18 years for the parent and teacher forms, 12–18 years for the self-report form
- **Reading level:** grade 3.7
- **Psychometric properties:** excellent normative sample, reliability and validity
- **Computerized reports:** narrative reports include reporting of all scores, comparisons between raters, intervention suggestions, and significance of changes in scores over time
- **Available Languages:** English and Spanish

Administration and Scoring Options

All of the forms can be administered via paper-and-pencil or in the MHS Online Assessment Center and scored via paper-and-pencil, the CEFI Scoring Software, or the MHS Online Assessment Center.

Report Options

CEFI reports can be obtained by using the Scoring Software or the MHS Online Assessment Center. There are three report types:

1. The ***Interpretive Report*** provides information about a single administration.
2. The ***Progress Monitoring and Treatment Effectiveness Report*** combines the results from up to four ratings by the same rater to examine changes in behavior that may have occurred over time.
3. The ***Comparative Report*** provides a multi-rater perspective by combining results from up to five different raters.

Normative Samples

The CEFI normative samples (parent *N* = 1,400, teacher *N* = 1,400, self-report *N* = 700) are representative of the U.S. population of children and adolescents. The samples were collected using a stratified sampling plan (based on the 2009 U.S. Census) based on age, gender, race/ethnicity, region, and parental education level. All representation of the sub-samples fell within 2% of the U.S. population targets (see Tables 1 to 4). Additionally, the normative samples included ratings of children who had a clinical diagnosis, or were eligible to receive special educational services according to IDEA criteria (parent = 10.9%, teacher = 12.7%, self-report = 9.7%).

Table 1. Age by Gender Distribution: CEFI Normative Samples

	Parent			Teacher			Self-Report		
Age	Male	Female	Total	Male	Female	Total	Male	Female	Total
5	50	50	**100**	50	50	**100**	-	-	**-**
6	50	50	**100**	50	50	**100**	-	-	**-**
7	50	50	**100**	50	50	**100**	-	-	**-**
8	50	50	**100**	50	50	**100**	-	-	**-**
9	50	50	**100**	50	50	**100**	-	-	**-**
10	50	50	**100**	50	50	**100**	-	-	**-**
11	50	50	**100**	50	50	**100**	-	-	**-**
12	50	50	**100**	50	50	**100**	50	50	**100**
13	50	50	**100**	50	50	**100**	50	50	**100**
14	50	50	**100**	50	50	**100**	50	50	**100**
15	50	50	**100**	50	50	**100**	50	50	**100**
16	50	50	**100**	50	50	**100**	50	50	**100**
17	50	50	**100**	50	50	**100**	50	50	**100**
18	50	50	**100**	50	50	**100**	50	50	**100**
Total	**700**	**700**	**1,400**	**700**	**700**	**1,400**	**350**	**350**	**700**

Table 2. Race/Ethnic Distribution: CEFI Normative Samples

Race/Ethnicity	Parent (%)	Teacher (%)	Self-Report (%)	U.S. Population (%)
Hispanic	22.0	22.0	22.0	**21.1**
Asian	4.0	3.8	4.0	**4.2**
Black	14.0	14.0	14.0	**13.9**
White	56.0	56.5	56.0	**56.5**
Other	4.0	3.7	4.0	**4.2**

Table 3. U.S. Region Distribution: CEFI Normative Samples

Region	Parent (%)	Teacher (%)	Self-Report (%)	U.S. Population (%)
Northeast	16.0	16.1	16.0	**17.0**
Midwest	22.1	22.0	22.0	**21.7**
South	37.9	37.9	38.0	**37.2**
West	24.1	24.0	24.0	**24.1**

Table 4. Parental Education Level Distribution: CEFI Normative Samples

Parental Education Level	Parent (%)	Self-Report (%)	U.S. Population (%)
No high school diploma	14.1	13.9	**14.7**
High school diploma/GED	27.9	28.0	**28.5**
Some college or associate's degree	29.9	30.0	**28.9**
Bachelor's degree	18.0	18.1	**17.6**
Graduate or professional degree	10.1	10.0	**10.3**

Reliability & Validity

Internal Reliability

The internal reliability coefficients (Cronbach's alpha) of the CEFI Full Scale are very high (range = .97 to .99) as shown in Table 5. Reliabilities for the nine scales were also excellent for parent (range = .84 to .93) and teacher (range = .90 to .96) raters. Self-report internal reliabilities are lower but still sufficiently high (range = .74 to .86).

Table 5. Internal Reliability Coefficients (Cronbach's Alpha)

Scale	Number of Items	Parent		Teacher		Self-Report
		5–11 Years	12–18 Years	5–11 Years	12–18 Years	12–18 Years
Full Scale	90	.98	.99	.99	.99	.97
Attention	12	.92	.93	.96	.96	.86
Emotion Regulation	9	.88	.90	.93	.93	.78
Flexibility	7	.84	.85	.90	.90	.77
Inhibitory Control	10	.89	.90	.94	.94	.80
Initiation	10	.88	.90	.92	.93	.80
Organization	10	.89	.92	.93	.94	.85
Planning	11	.91	.93	.95	.96	.85
Self-Monitoring	10	.85	.89	.91	.92	.78
Working Memory	11	.88	.89	.94	.94	.83

Test-Retest Reliability

Test-retest reliability was assessed over a 1- to 4-week interval by obtaining the correlation between CEFI standard scores from a general population sample. The test-retest values presented in Table 6 have been corrected for range instability in both distributions (see Guilford & Fruchter, 1978) and indicate excellent test-retest reliability.

Table 6. Test-Retest Reliability Coefficients (Pearson's *r*)

Scale	Parent (*N* = 171)	Teacher (*N* = 196)	Self-Report (*N* = 200)
Full Scale	.91	.90	.77
Attention	.88	.88	.74
Emotion Regulation	.87	.85	.74
Flexibility	.80	.82	.86
Inhibitory Control	.88	.86	.79
Initiation	.87	.86	.79
Organization	.89	.88	.86
Planning	.87	.89	.82
Self-Monitoring	.84	.89	.74
Working Memory	.89	.91	.79

Note. All *r*s significant, $p < .001$. All correlations were corrected for range instability.

Inter-Rater Reliability & Consistency Between Rater-Types

Inter-rater reliability results indicate substantial levels of rater agreement were found across all CEFI Scales for parent raters (r = .73 to .88), and moderate correlations were found for teacher raters (r = .54 to .68), indicating good levels of inter-rater reliability (see Table 7).

The examination of the consistency between rater-types (i.e., parent, teacher, and self-report ratings of the same child) also provides a way to examine the construct validity of the CEFI because the parent, teacher, and self-report forms contain the same items. Although some degree of similarity is expected between raters, some inconsistency is expected because the different raters have different opinions about, and different observations of, the child's behavior. Correlation coefficients between parent and teacher, parent and self-report, and teacher and self-report forms are presented for a sample of 126 children and youth (see Table 7). A sufficient amount of consistency between responses on the parent and self-report forms, parent and teacher forms, and teacher and self-report forms of the same child was found, thereby providing support for the construct validity of the CEFI. The correlations suggest that while there is a good degree of consistency between ratings, results from parents, teachers, and self-reports can differ. This underscores the importance of obtaining information from multiple informants.

Table 7. Inter-Rater Reliability and Consistency Between Rater-Types Coefficients (Pearson's *r*)

	Inter-Rater Reliability		Consistency Between Rater-Types		
Scale	Parent to Parent (N = 100)	Teachers to Teacher (N = 98)	Parent to Teacher (N = 126)	Parent to Self-Report (N = 126)	Teacher to Self-Report (N = 126)
Full Scale	.88	.68	.79	.71	.68
Attention	.86	.63	.76	.63	.64
Emotion Regulation	.73	.54	.58	.58	.39
Flexibility	.76	.63	.72	.44	.36
Inhibitory Control	.84	.64	.69	.65	.54
Initiation	.84	.57	.76	.67	.70
Organization	.86	.67	.73	.68	.61
Planning	.85	.68	.73	.63	.66
Self-Monitoring	.80	.68	.75	.56	.52
Working Memory	.82	.61	.72	.56	.51

Note. All *r*s significant, $p < .001$. All correlations were corrected for range instability.

Factor Structure

To determine the underlying structure of the CEFI, data from the normative samples were examined in a series of item-level and scale-level exploratory factor analyses. A series of procedures to evaluate the number of factors to retain (e.g., parallel analysis, scree plot tests, very simple solution criterion, ratio of first and second eigenvalues) were applied. Results indicated that a one factor solution best explained the data for both the item-level and scale-level factor analyses. Furthermore, this result was found to be very consistent across genders, age groups, race/ethnicities, and clinical statuses (coefficients of congruence ≥ .98 for all groups) indicating that the one factor solution of the CEFI was the same across these demographic groups. Taken as a whole, the data from this large study of nationally representative samples of individuals aged 5 to 18 years as rated by a parents, teachers, and youth self-reports indicate that the behaviors rated on the CEFI represent a single construct which can be interpreted as executive function.

Mean Score

Differences by Group

In order to test the criterion-related validity of the CEFI, data were collected from three samples of children and youth who were previously diagnosed with a clinical disorder of either ADHD (parent *N* = 172, teacher *N* = 142, self-report *N* = 118), an Autism Spectrum Disorder (ASD; parent *N* = 51, teacher *N* = 48), or a Mood Disorder (parent *N* = 37, teacher *N* = 30, self-report *N* = 28). It was expected that youth diagnosed with these clinical disorders should have executive function deficits, and therefore, should have lower CEFI Full Scale scores than youth from the general population.

CEFI standard scores for the three groups were compared to scores from matched samples of children and youth from the general population (see Figures 1 to 3). For the **ADHD** group analyses, all effects of group were significant for the parent, teacher, and self-report forms ($p < .001$), with moderate to large effect sizes (Cohen's d = -0.62 to -1.59). Results were statistically significant ($p < .001$) with large effect sizes for both the **ASD** analyses (Cohen's d = -0.99 to -1.41), and the **Mood Disorder** analyses (Cohen's d = -1.09 to -1.11). Children and youth from the three clinical groups (i.e., ADHD, ASD, and Mood Disorder) were expected to have executive function deficits when compared to general population samples, and the results of the analyses (i.e., moderate to large effects sizes across all forms for all three disorders) were consistent with these expectations. These results suggest that the CEFI is sensitive to differences in behaviors associated with executive function for these clinical groups.

Figure 1. Differences between ADHD and General Population Samples: CEFI Full Scale

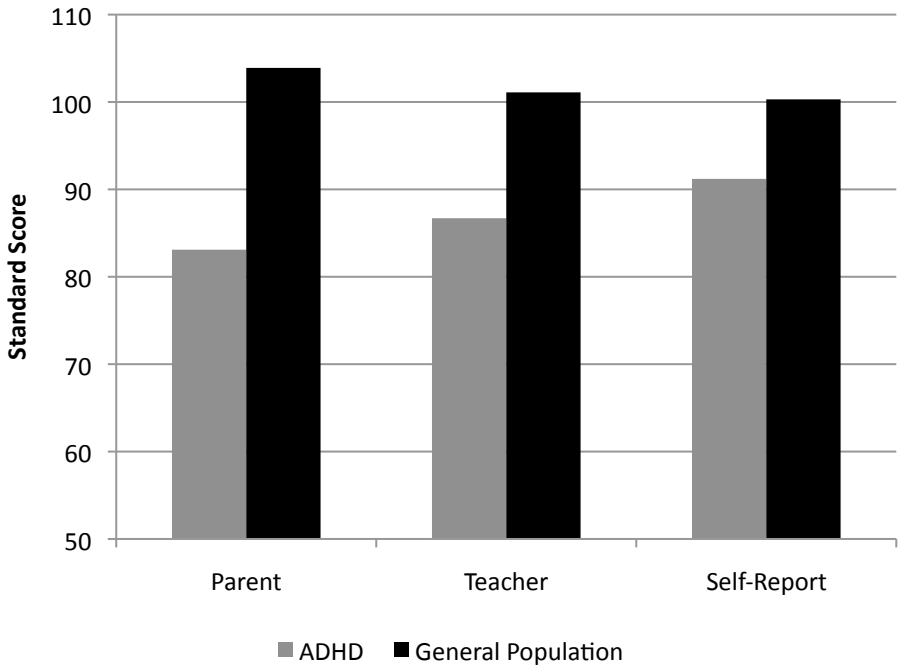


Figure 2. Differences between ASD and General Population Samples: CEFI Full Scale

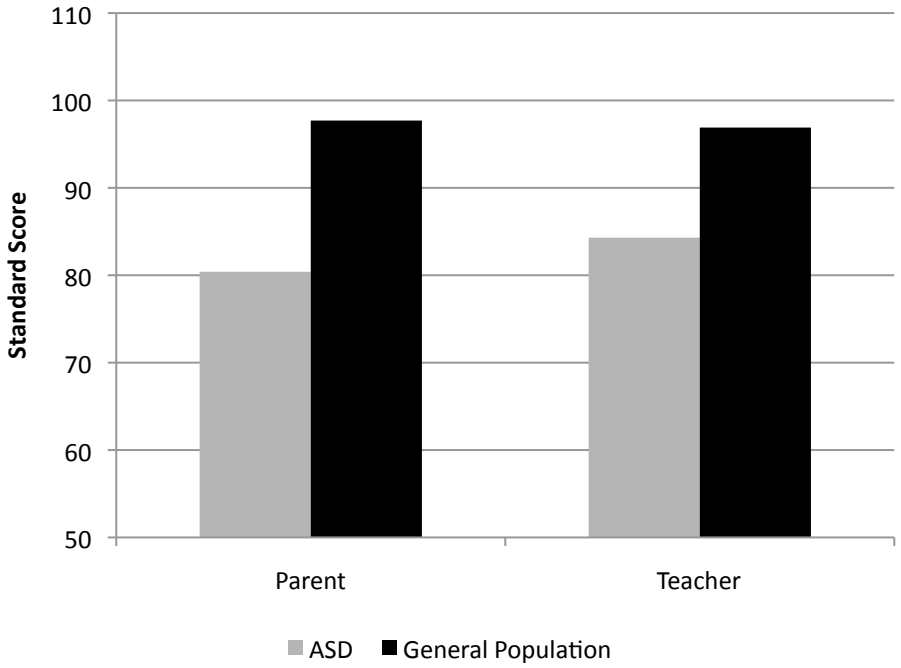
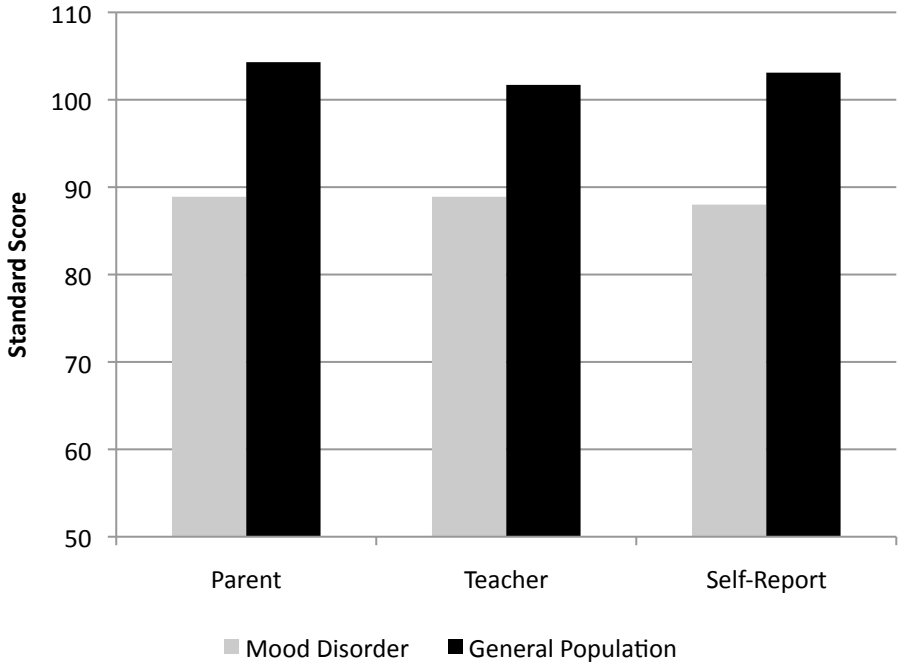


Figure 3. Differences between Mood Disorder and General Population Samples: CEFI Full Scale



Relationship Between the CEFI and the BRIEF

In order to examine the criterion-related validity of the CEFI, samples of parents (*N* = 57), teachers (*N* = 51), and youth (*N* = 32) completed the CEFI forms along with the Behavior Rating Inventory of Executive Functioning (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000; Guy, Isquith, & Gioia, 2004). All of the youth had been previously diagnosed with ADHD. The BRIEF yields *T*-scores which have a mean of 50 and *SD* of 10 and is scaled so that higher scores indicate poor executive function. To simplify the examination of the means and relationships between these two scales, the BRIEF scores were converted to a mean of 100 and *SD* of 15 and reversed so that low scores indicate poor executive function to match the metric and directionality of the CEFI.

The means and *SD*s of the CEFI Full Scale and the BRIEF Global Executive Composite (both in original and converted metrics) as well as corrected correlations are provided in Table 8. The correlations between the CEFI Full Scale and the BRIEF Global Executive Composite ranged from .68 (Self-Report ratings) to .85 (Parent ratings), which indicate that the two scales yield scores that are moderately correlated, but do show some differences. There are differences between the mean scores on the CEFI and the BRIEF scores. The results indicate that BRIEF standard scores are much lower than the CEFI for both parent and teacher raters and to a lesser extent, for youth self-report raters. These findings likely reflect the differences between the nationally normed (CEFI) and locally normed (BRIEF) scales. That is, the norms for the CEFI were based upon nationally representative sample stratified according to U.S. Census data for race/ethnicity, parental education levels (for Parent and Self-Report Forms), and region. In contrast, the standard scores for the BRIEF parent and teacher forms' were based on "normative data samples … obtained through public and private school recruitment in urban, suburban, and rural settings in the State of Maryland. A total of 25 schools were sampled" (Goia, Isquith, Guy, Kenworthy, 2000; p. 43). Differences between the CEFI and BRIEF mean scores (and the correlations between these scores) can reasonably be attributed to the manner in which the two scales were standardized and normed.

Table 8. Correlations between the CEFI Full Scale and the BRIEF Global Executive Composite

Form	*r*	*N*	CEFI Full Scale		BRIEF Global Executive Composite			
					Converted Scores		Original Scores	
			M	*SD*	*M*	*SD*	*M*	*SD*
Parent	.85	57	81.9	11.7	71.8	13.7	68.8	9.1
Teacher	.64	51	87.4	11.1	71.2	23.7	69.2	15.8
Self-Report	.68	32	90.2	14.2	86.7	15.9	58.8	10.6

Note. All correlations significant, $p < .01$. All correlations were corrected for range instability. CEFI Standard scores have a normative *M* = 100, *SD* = 15, with lower scores indicating more executive function problems. BRIEF scores were converted to have a normative *M* =100 and *SD* = 15 and scaled so that lower scores indicate more executive function problems (like the CEFI). Original BRIEF scores expressed as *T*-Scores (normative *M* = 50 and *SD* = 10) with higher scores indicating more executive function problems are also provided.

References

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

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

Guilford, J. P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th ed.). New York: McGraw-Hill.