# Behavior Intervention Monitoring Assessment System (BIMAS™)

# 10 Reliability

Reliability is defined as "the consistency of scores obtained by the same person when re-examined with the same test on different occasions, or with different sets of equivalent items, or under other variable examining conditions" (Anastasi, 1988, p. 102). This chapter presents results of reliability analyses of the Behavior Intervention Monitoring Assessment System (BIMAS™) Standard[1] (i.e., the BIMAS–Teacher [BIMAS–T], BIMAS–Parent [BIMAS–P], and BIMAS–Self-Report [BIMAS–SR]), including the results of internal consistency, standard error of measurement (*SEM*), test-retest reliability, and consistency between raters analyses.

## Overview of Results

Results of the reliability analyses revealed that the BIMAS forms have good levels of internal consistency, with Cronbach's alpha values from the total sample ranging from .81 to .91 on the BIMAS–T, from .77 to .90 on the BIMAS–P, and from .75 to .88 on the BIMAS–SR. Good levels of temporal stability (test-retest reliability) were found when the BIMAS was taken twice within a 2- to 4-week period (without any intervention), with correlation coefficients (Pearson's *r*) ranging from .85 to .91 on the BIMAS–T, from .79 to .96 on the BIMAS–P, and from .81 to .90 on the BIMAS–SR (all *r*s significant, $p < .001$). A good level of consistency between raters (i.e., teacher and self-report; parent and self-report; teacher and parent) was found on ratings of the same child with Pearson's *r* ranging from .54 to .86 across all scales (all *r*s significant, $p < .001$).

## Internal Consistency

One measure of a test's reliability is internal consistency, which is assessed using Cronbach's alpha (Cronbach, 1951). Cronbach's alpha ranges from 0.0 to 1.0, and is a function of the following parameters: "(a) the interrelatedness of the items in a test or scale, and (b) the length of the test" (John & Benet-Martinez, 2000, p. 343). There is no one level or cut-off of Cronbach's alpha that universally denotes satisfactory reliability. Adequate alpha depends on how many items are on the scale (the greater the number of items, the higher alpha tends to be), the scale's purpose, and the construct being measured. Internal consistency estimates (based on Cronbach's alpha) for the BIMAS were computed on a weighted sample comprising 85% normative cases and 15% clinical cases. For the internal consistency analyses, clinical cases were added to the normative sample to ensure an adequate level of variability in the data and to reflect real-world applications where there is a mix of youth with and without clinical diagnoses. Tables 10.1 to 10.3 present Cronbach's alpha values for this final weighted sample. The BIMAS–T, BIMAS–P, and BIMAS–SR were all found to demonstrate high levels of internal consistency for the majority of the scales. Specifically, for the total samples, BIMAS–T alpha values ranged from .81 to .91, BIMAS–P values ranged from .77 to .90, and BIMAS–SR values ranged from .75 to .88. With few exceptions, adequate to excellent levels of internal consistency were found across age and gender groups.

[1] Since only the norm-referenced standard form is discussed in this chapter, "BIMAS" is used to denote the BIMAS Standard throughout the chapter.

**Table 10.1. Cronbach's Alpha: BIMAS–T Standard**

| Gender and Age Group | | Behavioral Concern Scales | | | Adaptive Scales | |
|---|---|---|---|---|---|---|
| | | Conduct | Negative Affect | Cognitive/Attention | Social | Academic Functioning |
| **Total Sample** | | **.91** | **.85** | **.91** | **.85** | **.81** |
| Combined Gender | 5–6 | .85 | .79 | .91 | .76 | .66 |
| | 7–9 | .81 | .75 | .87 | .81 | .70 |
| | 10–11 | .86 | .84 | .91 | .84 | .80 |
| | 12–13 | .92 | .87 | .93 | .89 | .86 |
| | 14–16 | .93 | .87 | .92 | .88 | .81 |
| | 17–18 | .93 | .86 | .89 | .88 | .83 |
| Male | 5–6 | .87 | .75 | .92 | .83 | .58 |
| | 7–9 | .85 | .76 | .88 | .83 | .71 |
| | 10–11 | .87 | .82 | .90 | .83 | .71 |
| | 12–13 | .93 | .91 | .93 | .89 | .88 |
| | 14–16 | .94 | .89 | .92 | .89 | .76 |
| | 17–18 | .94 | .87 | .86 | .86 | .81 |
| Female | 5–6 | .81 | .84 | .87 | .58 | .72 |
| | 7–9 | .65 | .74 | .83 | .79 | .69 |
| | 10–11 | .84 | .86 | .91 | .84 | .87 |
| | 12–13 | .92 | .81 | .92 | .88 | .84 |
| | 14–16 | .92 | .84 | .93 | .86 | .86 |
| | 17–18 | .93 | .86 | .92 | .90 | .84 |

**Table 10.2. Cronbach's Alpha: BIMAS–P Standard**

| Gender and Age Group | | Behavioral Concern Scales | | | Adaptive Scales | |
|---|---|---|---|---|---|---|
| | | Conduct | Negative Affect | Cognitive/Attention | Social | Academic Functioning |
| **Total Sample** | | **.87** | **.82** | **.90** | **.84** | **.77** |
| Combined Gender | 5–6 | .79* | .69 | .90 | .78 | .49 |
| | 7–9 | .83 | .77 | .87 | .76 | .61 |
| | 10–11 | .82 | .80 | .90 | .75 | .71 |
| | 12–13 | .89 | .85 | .93 | .90 | .82 |
| | 14–16 | .90 | .86 | .89 | .86 | .84 |
| | 17–18 | .86 | .82 | .88 | .88 | .81 |
| Male | 5–6 | .78* | .59 | .91 | .82 | .49 |
| | 7–9 | .82 | .81 | .87 | .72 | .66 |
| | 10–11 | .86 | .85 | .90 | .71 | .61 |
| | 12–13 | .88 | .87 | .92 | .88 | .81 |
| | 14–16 | .90 | .86 | .90 | .81 | .83 |
| | 17–18 | .87 | .81 | .85 | .89 | .78 |
| Female | 5–6 | .80† | .76 | .87 | .69 | .47 |
| | 7–9 | .82 | .74 | .86 | .77 | .55 |
| | 10–11 | .73* | .67 | .88 | .82 | .81 |
| | 12–13 | .90 | .82 | .94 | .91 | .83 |
| | 14–16 | .91 | .86 | .86 | .91 | .84 |
| | 17–18 | .87 | .83 | .91 | .88 | .85 |

*Note.* *Calculated with item 32 (was suspected of smoking or chewing tobacco) taken out due to null variance. †Calculated with item 29 (was suspected of using alcohol and/or drugs) taken out due to null variance.

**Table 10.3. Cronbach's Alpha: BIMAS–SR Standard**

| Gender and Age Group | | Behavioral Concern Scales | | | Adaptive Scales | |
|---|---|---|---|---|---|---|
| | | Conduct | Negative Affect | Cognitive/Attention | Social | Academic Functioning |
| **Total Sample** | | **.88** | **.85** | **.87** | **.83** | **.75** |
| Combined Gender | 12–13 | .87 | .85 | .85 | .79 | .70 |
| | 14–16 | .88 | .85 | .87 | .83 | .75 |
| | 17–18 | .89 | .85 | .87 | .86 | .78 |
| Male | 12–13 | .88 | .83 | .84 | .79 | .71 |
| | 14–16 | .86 | .83 | .86 | .81 | .77 |
| | 17–18 | .89 | .83 | .86 | .87 | .74 |
| Female | 12–13 | .84 | .86 | .87 | .80 | .67 |
| | 14–16 | .90 | .86 | .88 | .84 | .72 |
| | 17–18 | .90 | .86 | .88 | .84 | .80 |

# Standard Error of Measurement

All measurements contain some error, which can be estimated with the standard error of measurement (*SEM*). To obtain a basic understanding of *SEM*, consider an individual's obtained score on the BIMAS as a reflection of the individual's "true" score on the test. Numerous factors may cause the obtained score to differ from, and fail to match exactly, with the individual's true BIMAS score. *SEM* provides an estimate of how much an individual's obtained score might vary from his/her true score. *SEM* is an estimate of the amount of error in the obtained scores. *SEM* values were calculated for each of the BIMAS scale *T*-scores using the *internal consistency* reliability estimates that were derived from the weighted normative and clinical samples (see *Internal Consistency* in this chapter). Tables 10.4 to 10.6 present *SEM* values for the BIMAS *T*-scores (see appendix H for *SEM* values for the BIMAS raw scores).

The *T*-score *SEM* values have been built into Confidence Intervals for the BIMAS *T*-scores (see *Confidence Intervals* in chapter 5, *Understanding and Interpreting BIMAS Scores*, for a description of how Confidence Intervals can be used in the interpretation process). Confidence Intervals surrounding the obtained *T*-score for every BIMAS scale are provided as an option in the computerized reports.

The Confidence Intervals were calculated by obtaining: (1) the standard error of measurement (*SEM*; see formula 1), (2) the Lower Bound of the Confidence Interval (see formula 2), and (3) the Upper Bound of the Confidence Interval (see formula 3).

**Formula 1.**

$$SEM = s_x\sqrt{1 - r_x}$$

where $s_x$ = the standard deviation of the *T*-scores earned by the weighted sample, and reliability = internal consistency estimate (Cronbach's alpha).

**Formula 2.**

*Lower Bound = Obtained Score - (z × SEM)*

where $z = 1.64$ for the 90% level of confidence, and $z = 1.96$ for the 95% level of confidence.

**Formula 3.**

*Upper Bound = Obtained Score + (z × SEM)*

where $z = 1.64$ for the 90% level of confidence, and $z = 1.96$ for the 95% level of confidence.

The *SEM* values were also used to determine the values needed for significance when comparing BIMAS results between raters (see *Statistically Significant Differences between Raters* in chapter 5, *Understanding and Interpreting BIMAS Scores*, for more information). For further information on the use and interpretation of the *SEM*, refer to McDonald (1999).

**Table 10.4. Standard Error of Measurement: BIMAS–T Standard *T*-Scores**

| Gender and Age Group | | Behavioral Concern Scales | | | Adaptive Scales | |
|---|---|---|---|---|---|---|
| | | Conduct | Negative Affect | Cognitive/ Attention | Social | Academic Functioning |
| **Total Sample** | | 2.58 | 3.94 | 3.42 | 4.20 | 4.36 |
| **Combined Gender** | 5–6 | 2.88 | 3.83 | 3.13 | 5.07 | 5.54 |
| | 7–9 | 3.47 | 4.30 | 3.69 | 4.34 | 5.29 |
| | 10–11 | 2.79 | 3.65 | 2.97 | 4.20 | 4.32 |
| | 12–13 | 2.27 | 3.76 | 3.23 | 3.78 | 3.70 |
| | 14–16 | 2.73 | 4.22 | 3.30 | 3.95 | 4.52 |
| | 17–18 | 2.49 | 4.07 | 3.92 | 3.81 | 4.02 |
| **Male** | 5–6 | 3.03 | 4.26 | 2.92 | 4.41 | 6.38 |
| | 7–9 | 3.10 | 4.19 | 3.59 | 4.16 | 5.33 |
| | 10–11 | 3.10 | 4.05 | 3.35 | 4.13 | 5.28 |
| | 12–13 | 2.32 | 3.27 | 3.32 | 3.89 | 3.45 |
| | 14–16 | 2.37 | 4.29 | 3.39 | 3.69 | 5.27 |
| | 17–18 | 2.36 | 3.83 | 4.23 | 4.28 | 4.06 |
| **Female** | 5–6 | 2.60 | 3.08 | 3.54 | 6.17 | 4.66 |
| | 7–9 | 4.32 | 4.64 | 3.99 | 4.47 | 5.12 |
| | 10–11 | 2.55 | 3.26 | 2.68 | 3.93 | 3.38 |
| | 12–13 | 2.21 | 4.33 | 3.10 | 3.82 | 3.92 |
| | 14–16 | 3.04 | 4.59 | 3.48 | 4.32 | 3.78 |
| | 17–18 | 2.87 | 4.26 | 3.39 | 3.46 | 4.14 |

**Table 10.5. Standard Error of Measurement: BIMAS–P Standard *T*-Scores**

| Gender and Age Group | | Behavioral Concern Scales | | | Adaptive Scales | |
|---|---|---|---|---|---|---|
| | | Conduct | Negative Affect | Cognitive/ Attention | Social | Academic Functioning |
| Total Sample | | 3.41 | 4.28 | 3.33 | 4.05 | 4.48 |
| Combined Gender | 5–6 | 4.02 | 5.22 | 3.46 | 4.53 | 6.41 |
| | 7–9 | 3.62 | 4.55 | 4.01 | 4.59 | 5.68 |
| | 10–11 | 3.73 | 4.37 | 3.30 | 4.93 | 4.97 |
| | 12–13 | 2.98 | 3.95 | 2.56 | 3.41 | 3.99 |
| | 14–16 | 2.99 | 3.82 | 3.32 | 3.92 | 3.58 |
| | 17–18 | 4.08 | 4.47 | 3.38 | 3.76 | 4.61 |
| Male | 5–6 | 4.14 | 5.87 | 3.37 | 4.27 | 6.54 |
| | 7–9 | 4.00 | 4.40 | 4.29 | 5.17 | 5.76 |
| | 10–11 | 3.39 | 3.80 | 3.25 | 4.99 | 5.93 |
| | 12–13 | 3.22 | 3.75 | 2.91 | 3.85 | 4.14 |
| | 14–16 | 3.14 | 3.78 | 3.04 | 4.72 | 3.82 |
| | 17–18 | 4.04 | 4.64 | 4.17 | 3.62 | 4.95 |
| Female | 5–6 | 3.79 | 4.34 | 3.74 | 5.08 | 6.25 |
| | 7–9 | 3.39 | 4.77 | 3.82 | 4.23 | 5.77 |
| | 10–11 | 4.33 | 5.30 | 3.47 | 4.12 | 3.94 |
| | 12–13 | 2.67 | 4.56 | 2.40 | 3.05 | 4.12 |
| | 14–16 | 2.88 | 3.93 | 3.86 | 3.06 | 3.75 |
| | 17–18 | 4.10 | 4.45 | 3.34 | 3.86 | 4.13 |

**Table 10.6. Standard Error of Measurement: BIMAS–SR Standard *T*-Scores**

| Gender and Age Group | | Behavioral Concern Scales | | | Adaptive Scales | |
|---|---|---|---|---|---|---|
| | | Conduct | Negative Affect | Cognitive/ Attention | Social | Academic Functioning |
| Total Sample | | 3.41 | 4.28 | 3.33 | 4.05 | 4.48 |
| Combined Gender | 12–13 | 3.28 | 3.94 | 4.02 | 4.62 | 4.88 |
| | 14–16 | 3.36 | 4.05 | 3.63 | 4.61 | 4.90 |
| | 17–18 | 3.33 | 4.07 | 3.76 | 4.13 | 4.16 |
| Male | 12–13 | 3.25 | 4.13 | 4.32 | 4.60 | 5.30 |
| | 14–16 | 3.58 | 4.36 | 3.58 | 4.45 | 4.67 |
| | 17–18 | 3.37 | 4.15 | 3.82 | 3.77 | 4.96 |
| Female | 12–13 | 3.08 | 4.16 | 3.76 | 3.95 | 5.00 |
| | 14–16 | 3.20 | 3.76 | 3.57 | 4.60 | 5.38 |
| | 17–18 | 3.29 | 3.90 | 3.69 | 4.80 | 3.59 |

# Test-Retest Reliability and Standard Error of Prediction

Test-retest reliability refers to the correlation of scores obtained from two separate administrations for the same youth by the same rater over a specified period of time. This type of reliability was assessed over a 2- to 4-week interval by obtaining *T*-score correlations for the BIMAS with a sample of 112 teachers, 83 parents, and 53 youth who completed the BIMAS twice (no interventions took place between the Time 1 and Time 2 administrations; see Table 10.7 for demographic characteristics of the test-retest samples). The correlations, as well as the means and standard deviations from Time 1 and Time 2 administrations, are provided in Tables 10.8 to 10.10. Across the three forms, $r$ ranged from .79 to .96 (all $p < .001$), indicating that the BIMAS has high test-retest reliability.

These test-retest values were then incorporated into the calculation of the standard error of prediction with the following formula:

$$\textit{Standard Error of Prediction} = S_x\sqrt{1 - r_{xy}}$$

where $S_x$ = the standard deviation of the normative sample, and $r_{xy}$ = test-retest reliability.

The standard error of prediction pertains to outcome assessment and is a method of determining how much scores may be expected to fluctuate over time due to random error when no intervention occurs between administrations. Table 10.11 presents standard error of prediction values for the BIMAS T-scores. These values provide an effective way to assess change over time (see *Scores for Progress and Outcome Monitoring* in chapter 5, *Understanding and Interpreting BIMAS Scores*).

**Table 10.7. Demographic Characteristics of the BIMAS Standard Test-Retest Reliability Samples**

| Demographic Characteristic of the Rated Youth | | Teacher | | Parent | | Self-Report | |
|---|---|---|---|---|---|---|---|
| | | *N* | % | *N* | % | *N* | % |
| Population | Non-Clinical | 112 | 100.0 | 83 | 100.0 | 53 | 100.0 |
| | Clinical | 0 | 0.0 | 0 | 0.0 | 0 | 0.0 |
| Gender | Male | 58 | 51.8 | 35 | 42.2 | 30 | 56.6 |
| | Female | 24 | 41.2 | 48 | 57.8 | 23 | 43.4 |
| Race/Ethnicity | Asian | 0 | 0.0 | 2 | 2.4 | 0 | 0.0 |
| | African American | 26 | 23.2 | 15 | 18.1 | 7 | 13.2 |
| | Hispanic | 17 | 15.2 | 15 | 18.1 | 20 | 37.7 |
| | White | 59 | 52.7 | 49 | 59.0 | 25 | 47.2 |
| | Other | 10 | 8.9 | 2 | 2.4 | 1 | 1.9 |
| **Total** | | **112** | **100.0** | **83** | **100.0** | **53** | **100.0** |
| **Age *M* (*SD*)** | | **7.14 (3.4)** | | **11.5 (3.7)** | | **15.4 (1.9)** | |
| **Days between administrations *M* (*SD*)** | | **21.6 (6.5)** | | **21.9 (7.2)** | | **20.2 (7.9)** | |

**Table 10.8. Test-Retest Reliability Coefficients: BIMAS–T Standard *T*-scores**

| Scale | | *r* | *N* | Time 1 | | Time 2 | |
|---|---|---|---|---|---|---|---|
| | | | | *M* | *SD* | *M* | *SD* |
| Behavioral Concern Scales | Conduct | .89 | 112 | 54.1 | 9.1 | 54.4 | 9.1 |
| | Negative Affect | .85 | 112 | 54.6 | 10.0 | 54.0 | 9.8 |
| | Cognitive/Attention | .91 | 112 | 51.8 | 13.0 | 51.7 | 14.0 |
| Adaptive Scales | Social | .91 | 112 | 46.1 | 13.0 | 46.3 | 13.0 |
| | Academic Functioning | .91 | 108 | 49.8 | 12.0 | 49.0 | 13.0 |

*Note.* All *r*s significant, $p < .001$. Sample sizes vary due to missing data.

**Table 10.9. Test-Retest Reliability Coefficients: BIMAS–P Standard *T*-scores**

| Scale | | *r* | *N* | Time 1 | | Time 2 | |
|---|---|---|---|---|---|---|---|
| | | | | *M* | *SD* | *M* | *SD* |
| Behavioral Concern Scales | Conduct | .79 | 82 | 49.6 | 7.9 | 49.6 | 7.4 |
| | Negative Affect | .91 | 79 | 54.4 | 12.0 | 53.9 | 11.0 |
| | Cognitive/Attention | .84 | 83 | 52.1 | 10.0 | 52.2 | 11.0 |
| Adaptive Scales | Social | .96 | 79 | 43.7 | 16.0 | 44.1 | 16.0 |
| | Academic Functioning | .80 | 83 | 52.3 | 8.4 | 53.2 | 8.1 |

*Note.* All *r*s significant, $p < .001$. Sample sizes vary due to missing data.

**Table 10.10. Test-Retest Reliability Coefficients: BIMAS–SR Standard *T*-scores**

| Scale | | *r* | *N* | Time 1 | | Time 2 | |
|---|---|---|---|---|---|---|---|
| | | | | *M* | *SD* | *M* | *SD* |
| Behavioral Concern Scales | Conduct | .81 | 52 | 45.8 | 7.0 | 45.6 | 7.5 |
| | Negative Affect | .87 | 52 | 45.4 | 9.6 | 46.2 | 10.8 |
| | Cognitive/Attention | .82 | 52 | 44.7 | 9.8 | 44.7 | 9.4 |
| Adaptive Scales | Social | .90 | 52 | 54.5 | 10.6 | 53.6 | 12.1 |
| | Academic Functioning | .85 | 49 | 52.9 | 8.4 | 52.3 | 9.2 |

*Note.* All *r*s significant, $p < .001$. Sample sizes vary due to missing data.

**Table 10.11. Standard Error of Prediction Coefficients: BIMAS Standard *T*-scores**

| Scale | | Rater | | |
|---|---|---|---|---|
| | | Teacher | Parent | Self-Report |
| Behavioral Concern Scales | Conduct | 3.25 | 4.64 | 4.39 |
| | Negative Affect | 3.86 | 3.05 | 3.58 |
| | Cognitive/Attention | 2.94 | 4.00 | 4.22 |
| Adaptive Scales | Social | 3.01 | 2.10 | 3.17 |
| | Academic Functioning | 3.03 | 4.53 | 3.85 |

# Consistency between Raters

Because the BIMAS–T, BIMAS–P, and BIMAS–SR measure similar constructs, a degree of consistency is expected across rater types. The level of consistency was assessed by correlating scores from the different raters. Although *some* degree of similarity is expected between raters, it is nonetheless expected that a certain degree of incongruence will exist (i.e., the correlations should be moderate in size). This incongruence occurs because the various raters may have different opinions about, and different experiences with, the youth's behavior. This incongruence can also occur because the raters see the youth in different contexts. A sample of 162 youth who provided self-report ratings were also rated by a teacher and a parent (see Table 10.12 for a sample description). Correlation coefficients (Pearson's *r*) were calculated between each pair of raters (see Tables 10.13 to 10.15). As anticipated, the correlations were found to be moderate in size (all $p < .001$) for the teacher to self-report comparisons ($r = .54$ to $.69$) and the parent to self-report comparisons ($r = .59$ to $.69$). A great deal of consistency was found between teacher and parent ratings, with strong correlations between the two rater types on all scales ($r = .79$ to $.86$).

**Table 10.12. Demographic Characteristics of the BIMAS Standard Consistency Between Raters Sample**

| Demographic Characteristic of the Rated Youth | Group | *N* | % |
|---|---|---|---|
| **Population** | Non-Clinical | 70 | 43.2 |
| | Clinical | 92 | 56.8 |
| **Gender** | Male | 80 | 49.4 |
| | Female | 82 | 50.6 |
| **Race/Ethnicity** | Asian | 2 | 1.2 |
| | African American | 23 | 14.2 |
| | Hispanic | 22 | 13.6 |
| | White | 101 | 62.3 |
| | Other | 14 | 8.6 |
| **Total** | | **162** | **100.0** |
| **Age *M* (*SD*)** | | **13.9 (1.8)** | |

**Table 10.13. Consistency Between Rater *T*-scores: Teacher to Self-Report Ratings**

| Scale | | *r* | Teacher | | Self-Report | |
|---|---|---|---|---|---|---|
| | | | *M* | *SD* | *M* | *SD* |
| **Behavioral Concern Scales** | Conduct | .54 | 56.5 | 9.7 | 49.2 | 10.1 |
| | Negative Affect | .64 | 59.0 | 12.4 | 52.2 | 13.4 |
| | Cognitive/Attention | .69 | 54.7 | 12.9 | 48.6 | 10.0 |
| **Adaptive Scales** | Social | .59 | 41.0 | 11.8 | 45.0 | 8.0 |
| | Academic Functioning | .59 | 47.8 | 11.6 | 49.4 | 8.7 |

*Note.* $N = 162$. All *r*s significant, $p < .001$.

**Table 10.14. Consistency Between Rater *T*-scores: Parent to Self-Report Ratings**

| Scale | | *r* | Parent | | Self-Report | |
|---|---|---|---|---|---|---|
| | | | *M* | *SD* | *M* | *SD* |
| **Behavioral Concern Scales** | Conduct | .62 | 54.4 | 11.2 | 49.2 | 10.1 |
| | Negative Affect | .69 | 56.8 | 13.3 | 52.2 | 13.4 |
| | Cognitive/Attention | .67 | 53.8 | 10.4 | 48.6 | 10.0 |
| **Adaptive Scales** | Social | .59 | 44.2 | 10.4 | 45.0 | 8.0 |
| | Academic Functioning | .65 | 46.1 | 9.9 | 49.4 | 8.7 |

*Note.* $N = 162$. All *r*s significant, $p < .001$.

**Table 10.15. Consistency Between Rater *T*-scores: Teacher to Parent Ratings**

| Scale | | *r* | Teacher | | Parent | |
|---|---|---|---|---|---|---|
| | | | *M* | *SD* | *M* | *SD* |
| **Behavioral Concern Scales** | Conduct | .82 | 56.5 | 9.7 | 54.4 | 11.2 |
| | Negative Affect | .86 | 59.0 | 12.4 | 56.8 | 13.3 |
| | Cognitive/Attention | .84 | 54.7 | 12.9 | 53.8 | 10.4 |
| **Adaptive Scales** | Social | .79 | 41.0 | 11.8 | 44.2 | 10.4 |
| | Academic Functioning | .80 | 47.8 | 11.6 | 46.1 | 9.9 |

*Note.* $N = 162$. All *r*s significant, $p < .001$.