The Attention-Deficit Disorders Evaluation Scale—Second Edition, School Version (ADDES-2, SV) is a rating scale consisting of 60 items divided into two subscales: Inattentive (29 items) and Hyperactive–Impulsive (31 items). The test was designed to measure the three constructs of attention deficit hyperactivity disorder (ADHD) posited in The Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition: ADHD, combined type; ADHD, inattentive type; and ADHD, hyperactive–impulsive type. The ADDES-2, SV is administered individually; it takes about 20 minutes for an educator familiar with the target student to complete the scale. Appropriately used with students from 4 or 5 years to 18 years of age, the ADDES-2, SV is intended to accomplish the following (J. McCarney, 1995b, p. 5):

1. Screen for characteristics of ADHD
2. Provide a measure of ADHD
3. Contribute to the diagnosis of ADHD
4. Contribute to the development of program goals and objectives
5. Identify intervention activities for ADHD in educational environments

Each item is scored on a 5-point scale:

0 Does not engage in the behavior
1. One to several times per month
2. One to several times per week
3. One to several times per day
4. One to several times per hour

All items on all scales are to be completed; if an item is not appropriate for a particular student, it is scored as "Does not engage in the behavior."

**Scores**

Raw scores on each subscale can be converted to standard scores (mean = 10; standard deviation = 3). Five score-conversion tables are provided for females, based on their ages: 4 through 6, 7 through 8, 9 through 12, 13 through 15, and 16 through 18. Five score-conversion tables are also provided for males, but the age ranges are different: 4 through 7, 8 through 9, 10 through 13, 14 through 15, and 16 through 18. Subscale standard scores can be summed and converted to percentile ranks.

**Norms**

The ADDES-2, SV was normed on 5,795 students who ranged in age from 4–0 to 19 years. No age group has fewer than 329 students. In all, 30 states are represented in the norms. The author provides no information about his sampling plan or how students and teachers were recruited.
When compared with 1992 U.S. demographic data, the sample as a whole is generally representative in terms of gender, race, urban/rural residence, and mother’s occupation. The North Central region of the country is overrepresented, as are blue-collar workers. However, no information is provided about specific age groups—critical information for judging the adequacy of the scale's norm for the particular student being evaluated.

Reliability

Stability was estimated using 481 individuals randomly selected from the normative sample, who were rerated within 30 days of their initial rating. For total score and each subscale, an estimate of stability is provided for each age group and gender. All stability estimates exceed .90 for the total score and the Inattentive subscale. For the Hyperactive–Impulsive subscale, the stabilities for males all equal or exceed .90; three of the five stabilities for females equal or exceed .90, and the remaining two are in the high .80s.

Interrater agreement was evaluated by having 237 pairs of teachers rate 462 students selected at random from the standardization sample. The two ratings of each subject were correlated by ages (for boys and girls together). With one exception (age 16), all correlations were in the .80s.

Coefficient alpha was used to estimate the internal consistency of each subscale and the total score, based on the performance of the total normative sample.\(^1\) One

---

\(^1\) Although in the text the author says that KR-20 was used to estimate internal consistency, the table containing the values reports coefficient “alpha.” Because KR-20
Coefficient alpha (.98) is reported for each subscale and for the total score (.99). However, because this coefficient was calculated across a wide range of ages, internal consistency will be overestimated to the extent that age and ADDES-2, SV scores are correlated.

Validity

Overall, evidence for content validity appears good. Items for the ADDES-2, SV were derived from the literature and input from educational diagnosticians. Items were field-tested prior to standardization. Data from the standardization sample suggest that individual items produce distributions of responses and that individual items are associated with total scores.

Construct validity was investigated in two ways. First, although incompletely reported, factor analyses suggest that the ADDES-2, SV comprises as many as five factors. The first factor (inattention) is very strong, and the second factor (hyperactive–impulsive) is relatively weak. In a two-factor solution, 41 of the 60 scale items are correlated with both factors, and the correlation between subscales exceeds .7. Some evidence of construct validity is provided by a study comparing 72 male and 30 female students previously identified as having ADHD with 72 male and 30 female students drawn at random from the standardization sample. On average, both male and female students with ADHD earned substantially lower scores than the nondisabled students.

Can be used only with dichotomously scored data and the items in this scale are scored on a 5-point scale, we assume that “KR-20” is a typographical error.
To investigate the criterion-related validity, ADDES-2, SV ratings of students previously identified as having ADHD were compared with their ratings on five other scales: the Conners Teacher Rating Scale–28, the Conners Teacher Rating Scale–39, the ADHD Comprehensive Teacher's Rating Scale, the Children's Attention and Adjustment Survey (school form), and the Child Behavior Checklist (teacher report form). Although incompletely reported, the results of these studies appear to provide good support for concurrent validity. The ADDES-2, SV is highly correlated with subtests measuring inattention, impulsivity, hyperactivity, and so on, but it is essentially uncorrelated with other characteristics (such as withdrawal).

Summary

The ADDES-2, SV is a 60-item rating scale designed for use with students between 4 and 18 years of age. There are two subscales (Inattentive and Hyperactive–Impulsive). Separate norms are provided for boys and girls at five age groups. Although the overall sample appears to be generally representative of the United States, no information about the representativeness of the individual age–gender norm groups is presented. Scores tend to be stable, but interrater agreement is sufficient only for screening purposes. Data on internal consistency are inadequately reported. Content and criterion-related validity appear strong. Evidence of construct validity is quite limited.