Walker–McConnell Scale of Social Competence and School Adjustment (W-M)

The Walker–McConnell Scale of Social Competence and School Adjustment (H. M. Walker & McConnell, 1988) is a norm-referenced rating scale of social skills for use with elementary-age children. The instrument consists of three subscales (Teacher-Preferred Social Behavior, Peer-Preferred Social Behavior, and School Adjustment Behavior) intended to be used for "screening and identification of social skills deficits" (H. M. Walker & McConnell, 1988, p. 1). The W-M was "not designed as either a diagnostic or classification instrument" (p. 3). Information on students' social competence and school adjustment that can be used as part of referral or child-study team processes, as well as for program planning, is obtained from item subscale and total scores and through normative comparisons. The authors suggest that in addition to the national norms provided in the manual, a classroom normative reference can be gained by assessing typical peers of the target student on the W-M. This procedure also helps validate the ratings of the target student.

The W-M consists of 43 positively worded descriptions of social skills. Subscale 1 (Teacher-Preferred Behavior) consists of 16 items that assess sensitivity, empathy, cooperation, self-control, and maturity. Examples of items include "Shows empathy" and "Cooperates with peers." Subscale 2 (Peer-Preferred Behavior) has 17 items that address peer values and relations in social situations. Items include "Invites peers to play or share activities" and "Compromises when the situation calls for it." Subscale 3 (School Adjustment Behavior) has 10 items that assess competencies in academic settings. Sample items include "Displays independent study skills" and "Listens carefully to
teacher directions." All items are rated on a 5-point Likert scale from "never occurs" to "frequently occurs." The upper anchor point is defined in somewhat vague terminology ("frequently occurs" is defined as "a high rate"), and a rationale for the scaling choice is not provided. More explicit definitions may help raters make more objective evaluations of student behavior.

Scores

The W-M yields raw scores (subscale and total) by adding the numerical Likert ratings. Raw scores are converted to standard scores for the subscales (mean = 10; standard deviation = 3) and for the total scale (mean = 100; standard deviation = 15). Standard scores are converted to percentile rankings.

Norms

Normative data for the W-M are based on a sample of 1,812 students from grades K to 6. Sixty percent of these students were in grades 1, 2, and 4, but there were at least 125 students per grade. Data were collected over a two-year period from 1985 to 1987. Students were from 15 states, with 3 states (Alabama, Montana, and Oregon) supplying 48 percent of the sample. The authors provide data to show that the mean scores of subjects from the four geographic regions were all within one SEM. Other demographics (such as urban/suburban and ethnicity) appear representative of the nation.
Reliability

The authors supply three types of reliability data: test–retest, interrater, and internal consistency. In addition, item-total correlations (which we consider to be reliability data) are found in the manual's section on item validity.

Test–retest reliability is supported by data from several studies, some of which were conducted by other investigators. Coefficients range from .67 to .97 for studies with test–retest intervals of two to four weeks. A study of test–retest stability with antisocial and normal boys over a six-month period produced slightly lower correlations.

Interrater reliability studies produced what the authors call "modest agreement levels" (p. 33). What is missing is an estimate of interrater reliability using two same-type raters (for example, two teachers or two teacher's aides). Such data provide an estimate of the consistency of assessments of student functioning from a common frame of reference.

Internal-consistency data are presented in the form of alpha coefficients for the total norm sample. All subscales were above .95, and all coefficients by grade level were above .94. Collectively, these data support the claim that the W-M has adequate reliability for its intended purposes.

Validity

Twenty-four pages of the 50-page manual for the W-M are devoted to text, tables, and figures on validity data (including the item–total correlations discussed previously). The
authors report their own studies and those of other investigators using the W-M in research.

**Content Validity** The authors of the W-M examined the professional literature, existing teacher-rating instruments, and materials from other national projects to develop an initial item pool of 100 descriptors of school-related social skills. Then 83 items were selected for testing, and the final 43 items were chosen based on (a) item means, (b) item variances, (c) item–total correlations, and (d) item loadings from first- and second-order factor analyses. Clearly, the authors have employed empirical processes for the identification of the items included in the W-M.

**Criterion-Related and Discriminative Validity** Several studies are presented to support the criterion-related validity of the W-M. These are mostly supportive of concurrent validity. In a study using the Walker Problem Behavior Identification Checklist (WPBIC) as the criterion, 13 elementary students who had been referred for resource room services and 17 nondisabled students who had been referred for counseling were rated by teachers and by their parents. The WPBIC and W-M are scaled in opposite ways; thus the resulting validity coefficients were negative and significant (range = -.69 to -.89).

In another study using teacher ratings of social adjustment, peer sociometrics, and academic achievement variables as criteria, 65 elementary students were also rated on the W-M. Moderate to high correlations were found between the teacher ratings and the W-M subscale and total scores. Low to moderate correlations were found for peer sociometrics and achievement data. The W-M does not appear to correlate with measures of academic engagement.
Correlations with the Social Skills Rating System for Teachers (SSRS-T) provide potentially strong evidence of concurrent validity. The SSRS-T is an earlier version of the teacher scale of the Social Skills Rating System (Gresham & Elliott, 1990).

Unfortunately, the authors do not report the number of subjects or population(s) involved in the validity study. The total scores of the W-M and the SSRS-T correlated. The subscales of the W-M and factors of the SSRS-T correlated in the expected directions.

Discriminative validity of the W-M is supported by studies designed to document the sensitivity of the instrument in discriminating groups defined by researchers or based on school district classifications. Other discriminative-validity studies involved research-defined groups of antisocial and normal students, students categorized under a two-dimensional model of behavior (disciplinary problems and peer acceptance), sociometrically defined groups (for example, popular and rejected), and school-identified groups of at-risk students and students with disabilities. In most cases, the W-M appears sensitive to group differences and predicts membership or performance in expected directions. Scores on Subscales 1 and 2 and total score on the W-M appear to discriminate better than Subscale 3 scores.

Standard scores between 7 and 13 on the W-M are within "normal" limits. In plotting the "characteristic profiles" of five student groups (normal, learning disabled, residential severely emotionally disturbed, resource room, antisocial), the normal group scored highest, and the scores of the other four groups were lower. However, there is very little differentiation among these four groups, and, although the authors say these students show moderate to severe deficits on the W-M, nearly all subscale standard scores of these
groups are within one standard deviation of the mean. The W-M, as the authors state, is not to be used for differential diagnoses among disability groups.

**Construct Validity** The authors report evidence of construct validity and factorial validity. Subscale 2 is specifically supported by longitudinal research on 100 boys at risk for antisocial behavior. The construct validity of the W-M is also argued on the basis of data showing near-zero correlations with student age and very low correlations with student gender, although gender differences favoring females might be expected.

**Summary**

The Walker–McConnell Scale of Social Competence and School Adjustment is a 43-item rating scale for screening and identification of social-skills deficits in elementary school students. The W-M is easy to administer and score; it yields raw scores, standard scores, and percentiles. The norms for the W-M are adequate, although there is some overrepresentation of western states. Reliability data to support its use include test–retest, interrater, and internal-consistency coefficients. Coefficients are generally above .80. Content, criterion-related, and construct validity data are convincing for the intended purposes of the device. Though three subscales are identified, the W-M is best used for normative comparisons as a global measure of social competence. Subscales and items within subscales may provide teachers with indications of specific areas of skills deficits for individual students. Educators must remember, however, that item-reliability data are not provided, and the authors' suggestions of remediation based on individual item data need further research support.