Original Research

Hammill Institute on Disabilities

Assessment for Effective Intervention
1–16

DOI: 10.1177/1534508418808598
aei.sagepub.com
SAGE

Psychometric Fundamentals of the Social Skills Improvement System: Social–Emotional Learning Edition Rating Forms

Frank Gresham, PhD[1], Stephen Elliott, PhD[2], Sarah Metallo, PhD[1], Shelby Byrd, MS[1], Elizabeth Wilson, MA[1], Megan Erickson, MA[1], Kaitlin Cassidy, MA[1], and Robert Altman, PhD[3]

Abstract

This study described the development of the *Social Skills Improvement System Social Emotional Learning Edition Rating Forms* (SSIS SEL RF) for teachers, parents, and students. This new multirater assessment is a reconfiguration of the SSIS Rating Scales items inspired by the CASEL Social Emotional Competency framework. The internal structure and score reliability estimates were examined across three raters for a common sample of more than 200 individual children ages 3 to 18 years. Confirmatory factor analyses tested against the CASEL five-dimensional SEL theoretical model demonstrated adequate fit for the SSIS SEL Parent and Student RFs and mediocre fit of the Teacher RF. Internal consistency, test-retest, and interrater reliability estimates for scores on each of the SSIS SEL RFs all met or exceeded acceptable criteria. Thus, researchers and practitioners interested in measuring the social–emotional behavior of children ages 3 to 18 can expect reliable scores and structurally meaningful behavior content within the Collaborative on Academic Social Emotional Learning (CASEL) SEL competency framework. Limitations to the present findings and suggestions for future research conclude the report.

Keywords

SSIS SEL edition rating forms, CASEL SEL model, confirmatory factor evidence, score reliability estimates

Over the past decade, there has been a substantial interest nationally and internationally in children's social–emotional learning (SEL) that has resulted in calls to increase educational services to improve children's SEL skills (Collaborative on Academic Social Emotional Learning [CASEL], 2015; Organisation for Economic Co-operation and Development, 2015; Weissberg & Cascarino, 2013; World Economic Forum, 2016). This concern about children's well-being is well-founded because many children continue to be at risk for adjustment problems in areas of educational, psychosocial, and vocational spheres of functioning (Taylor, Oberle, Durlak, & Weissberg, 2017). Difficulties in social–emotional functioning are characteristic of individuals with a range of disabilities, including emotional and behavioral disorders (Gresham, Cook, Crews, & Kern, 2004; Maag, 2005; Walker & Gresham, 2014), specific learning disabilities (Gresham, MacMillan, Bocian, Ward, & Forness, 1998), attention-deficit/hyperactivity disorder (MTA Cooperative Group, 1999), conduct disorder (Conduct Problems Prevention Research Group, 1999, 2002), mild intellectual disability (Gresham & Reschly, 1987), and autism spectrum disorders. In summary, as many as one third of all children, especially children with disabilities, do not adequately develop key SEL skills known to enable academic achievement and success in the workplace (CASEL, 2015; DiPerna, Volpe, & Elliott, 2002, 2005).

The importance of children's social competence has become increasingly clear across a variety of settings and outcomes. Social competence has been variously termed as *social skills, soft skills, personal adjustment, social functioning, social behavior*, and, perhaps most frequently of late, *social–emotional learning*. These terms evolved from different eras and different disciplines; however, they appear to be referring to the same, or very similar, core constructs. It is clear that children

[1]Louisiana State University, Baton Rouge, USA
[2]Arizona State University, Tempe, USA
[3]A-PsychEd Publication Services, Minneapolis, MN, USA

Corresponding Author:
Frank M. Gresham, Louisiana State University, 119 Audubon Hall, Baton Rouge, LA 70803, USA.
Email: frankmgresham@gmail.com

and youth with stronger SEL skills and higher functioning levels tend to have greater academic success, tend to be more socially adjusted, and have lower risk of serious adulthood psychopathology (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Greenberg et al., 2003; Weissberg, Durlak, Domitrovich, & Gullotta, 2015).

CASEL Model of SEL Competencies and the Assessment of SEL Skills

SEL has been defined as the process of acquiring knowledge, skills, attitudes, and beliefs to identify and manage emotions; to care about others, to make good decisions, to behave ethically and responsibly, to develop positive relationships, and to avoid negative behaviors (CASEL, 2013). Based on this definition, the Collaborative on Academic Social Emotional Learning (CASEL) posited a theoretical model of SEL, often referred to as the CASEL Five (CASEL, 2013, 2015). The CASEL Five are defined as follows:

- *Self-Awareness Skills*, defined as the ability to accurately recognize one's emotions and thoughts and their influence on behavior. This includes accurately assessing one's strengths and limitations and possessing a well-grounded sense of confidence and optimism.
- *Self-Management Skills*, defined as the ability to regulate one's emotions, thoughts, and behaviors effectively in different situations. This includes managing stress, controlling impulses, motivating oneself, and setting and working on achieving personal and academic goals.
- *Social Awareness Skills*, defined as the ability to take the perspective of and empathize with others from diverse backgrounds and cultures, to understand social and ethical norms for behavior, and to recognize family, school, and community resources and supports.
- *Relationship Skills*, defined as the ability to establish and maintain healthy and rewarding relationships with diverse individuals and groups. This includes communicating clearly, listening actively, cooperating, resisting inappropriate social pressure, negotiating conflict constructively, and seeking and offering help when needed.
- *Responsible Decision-Making Skills*, defined as the ability to make constructive and respectful choices about personal behavior and social interactions based on consideration of ethical standards, safety concerns, social norms, the realistic evaluation of consequences of various actions, and the well-being of self and others (CASEL, 2013, p. 9).

Interestingly, many of the specific skills described in the CASEL model traditionally have been part of the earlier versions of the *Social Skills Rating System* (Gresham & Elliott, 1990) and the *SSIS Rating Scales* (Gresham & Elliott, 2008), two widely used social skills assessments. Numerous definitions of social skills exist, and nearly all describe behaviors that facilitate the initiation and maintenance of positive social relationships, contribute to peer acceptance, allow individuals to cope with and adapt to the demands of the social environment, and result in satisfactory school adjustment while enabling academic achievement (Gresham, 2018). These behaviors are part of a number of well-regarded assessments and clearly are embedded within many, if not all, the SEL core competencies advanced by CASEL. Therefore, historically labeled as social skills, the vast majority of these skills also represent the core SEL competency domains in the CASEL model.

The CASEL Five model (CASEL, 2013) has been influential in the development of dozens of school-based intervention programs in the United States and in the national curricula of countries such as Australia, Great Britain, Northern Ireland, Singapore, and Sweden (Humphrey, Lendrum, Wigelsworth, & Greenberg, 2016). It has not, until recently, however, directly influenced the development of assessments of social, emotional, and academic skills commonly targeted within these programs (Elliott, Frey, & Davies, 2015). An assessment guide of assessments of SEL skills is forthcoming in 2018 by an independent panel at CASEL (visit CASEL.org for information). Suffice it to say, there has not been a comprehensive norm-referenced assessment of SEL skills fully aligned with the popular CASEL model. In the remainder of this article, we examine the psychometric characteristics of a multi-informant (parent, teacher, and student) rating form that is based on the SSIS Rating Scale and designed to measure the SEL skills identified by CASEL.

Extant SEL Assessment Tools

Two major reviews of measures of social and emotional skills for children and youth have been published (i.e., Crowe, Beauchamp, Catroppa, & Anderson, 2011; Humphrey et al., 2011). Many of the measures reviewed focus on social–emotional problem behaviors rather than positive or strength-focused behaviors advanced by current SEL competency models such as CASELs. From the dozens of assessments that purport to measure SEL competencies, four besides the new SSIS SEL assessments are technically sound and appear to measure a number of positive SEL skills. These are the *Behavioral and Emotional Rating Scale–Second Edition* (BERS-2; Epstein, 2004), *Devereux Student Strengths Assessment* (DESSA; LeBuffe, Shapiro, & Naglieri, 2009/2014), *Social-Emotional Assets and Resilience Scales* (SEARS: Merrell, 2011), and the *Social Skills Improvement System–Rating Scales* (SSIS-RS: Gresham & Elliott, 2008). Each of these assessments is described next.

BERS-2

The BERS-2 measures the personal strengths and competencies of children ages 5 years 0 months to 18 years 11 months and is a multimodal assessment system (i.e., child self-ratings, parent ratings, and teacher ratings) intended to help in the identification of students with emotional and behavioral difficulties. The BERS-2 rating scales can each be completed in approximately 10 min and measure six aspects of children's strengths: interpersonal, involvement with family, intrapersonal, school functioning, affective, and career strengths. The BERS-2 scales were normed on representative samples of children without disabilities; the Teacher Scale was also normed on children with emotional and behavioral problems. The internal consistency reliability of the BERS-2 subtests all exceeded .80, and the overall score was .95. The Examiner's Manual reports on numerous studies that provide evidence to support the BERS's content, construct, and criterion-related validity (e.g., Epstein, 1999; Harniss, Epstein, Ryser, & Pearson, 1999). This measure was developed prior to the formulation of the CASEL SEL Competence framework. Although some of its items clearly relate to the five competency domains, their alignment has not been formalized nor has it been a common measure reported in the SEL program screening or outcome evaluation studies.

DESSA

The DESSA is a 72-item, standardized, norm-referenced behavior rating scale that assesses 8 social and emotional competence subscales (i.e., Personal Responsibility, Optimistic Thinking, Goal-Directed Behavior, Social Awareness, Decision-Making, Relationship Skills, Self-Awareness, and Self-Management) that serve as protective factors for children kindergarten to eighth grade. This assessment is completed by parents and teachers; there is no student/child self-rating scale. As described by Shapiro, Kim, Robitaille, and LeBuffe (2017), the DESSA was standardized on a sample of nearly 2,500 students. There are not separate norms for age groups or for boys and girls. The DESSA Technical Manual and User's Guide (2009) reports very high internal reliability for parents' and teachers' Social Emotional Composite (SEC) ratings with the various subscales ranging from .82 to .94. Test-retest reliability for the SEC is reported as .90 for parents and .94 for teachers. Other researchers (Shapiro, Accomazzo, & Robitalle, 2017) have reported high levels (88%) of cross-informant agreement between teachers and staff. Finally, a number of studies provide solid evidence for the content and construct validity of the DESSA (e.g., Nickerson & Fishman, 2009).

SEARS

The SEARS is a multi-informant, strengths-based, social–emotional assessment system for assessing children and adolescents who exhibit a variety of clinical problems or who are at high risk for developing such problems. It provides four rating forms: Parent (39 items), Teacher (41 items), Child (35 items, Grades 3–6), and Adolescent (35 items, Grades 7–12). According to the author, all forms measure the common constructs of self-regulation, responsibility, social competence, and empathy. Items representing these constructs are similar but not always the same across forms. These forms were normed on a large, but not nationally representative, sample of children in kindergarten through Grade 12. There is good evidence that the internal consistency reliability is greater than .90 for each form and test-retest reliability was good but in the .67 to .84 range. There is also reasonable convergent validity evidence when compared with other measures of social skills and life satisfaction (Nese et al., 2012). This measure also was developed prior to the formulation of the CASEL SEL Competence framework, and although some of its items clearly relate to the five competency domains, their alignment has not been formalized nor has it been a frequently used measure in research or practice.

SSIS-RS

The SSIS-RS (and its predecessor, the *Social Skills Rating System* [SSRS]; Gresham & Elliott, 1990) have a 3-decade history of providing technically sound assessments of children's and youths' social skills (Gresham, Elliott, Vance, & Cook, 2011). The SSRS and SSIS-RS assessments have been recognized by many researchers as two of the technically best social–emotional assessments (e.g., Humphrey et al., 2011). The SSIS-RS (Gresham & Elliott, 2008) is a broad-band, multirater (teacher, parent, and student) assessment of students' social behavior that examines teacher–student relations, peer interactions, and academic performance (teacher rating scale only). Item development of the SSIS-RS was based on a broad review of the empirical literature on social skills deficits in special populations, reviews of published empirical studies using SSRS, and research on the relationship between specific social behaviors and important social outcomes for children and youth. The SSIS-RS solicits this information from three rating sources (teachers, parent, and students) in Grades 3 through 12 and from parents and teachers for children ages 3 to 5 years. The SSIS-RS measures seven social skills subscales (Communication, Cooperation, Assertion, Responsibility, Empathy, Engagement, and Self-Control) generated empirically. The teacher and parent forms include problem behaviors

from five subdomains (Externalizing, Bullying, Hyperactivity/Inattention, Internalizing, and Autism Spectrum) with the Bullying and Autism Spectrum subscales. The teacher rating scale continued to include an Academic Competence scale measuring student performance in reading, mathematics, motivation, parental support, and general cognitive functioning. The SSIS-RS was normed in 2006–2007 on a nationwide sample totaling 4,700 children and adolescents ages 3 through 18 years who were assessed at 115 sites in 36 states. Demographic targets for the norm sample were based on *Current Population Survey, March 2006* (U.S. Census Bureau, 2006) and were applied to the three norm groups (ages 3–5 years, 5–12 years, and 13–18 years). The SSIS-RS Manual provides extensive validity evidence based on test content, internal structure, intercorrelations among scales and subscales, item-total correlations, and relations with other variables (Gresham & Elliott, 2008). Intercorrelations among scales and subscales for each form are moderate to high for the social skills and problem behavior scales. Item-total correlations across forms by age tend to be moderate to high, many of which exceed .70 to .80. Correlations between the SSIS-RS and the *Behavioral Assessment System for Children, Second Edition* (BASC-2) are moderate to high, depending on the scales and subscales. For example, the median correlations between the SSIS-RS total social skills score and the teacher form of the BASC-2 social skills score are .78 and .69 for the teacher and parent forms, respectively. Correlations between the SSIS-RS total social skills scores and the socialization scores of the *Vineland Adaptive Behavior Scales, Second Edition* (Vineland II) are .65 and .44 for the teacher and parent forms, respectively.

Purpose and Expected Outcomes of the Present Study

To advance educational practices and corresponding research on SEL skills, sound assessments are needed for the purpose of identifying children's SEL skills in need of improvement and evaluating outcomes of programs designed to improve these SEL competencies. As reviewed, there are four extant rating scales with good to excellent psychometric qualities that focus on children's and adolescents' positive social emotional skills. Yet, only one of these assessments, the DESSA, purports to measure the popular model of SEL skills theorized as highly relevant to social and academic success. The DESSA, however, has some limitations in that it does not offer a self-rating form, is not normed with high school-age students, and is not well aligned directly with any SEL intervention programs. As such, the current study sought to evaluate the psychometric properties (reliability and structural validity evidence) of the scores from the SSIS SEL RF (Gresham & Elliott, 2017), an assessment that addresses each of these limitations. The SSIS SEL RF items are those of the original SSIS Rating Scales for Teachers, Parents, and Students (Gresham & Elliott, 2008). Specifically, in this report, we provide an examination of the internal structure of the SSIS SEL RF. We expected to fit adequately within the five competencies of the CASEL SEL competency framework. To further explore this expectation, we conducted a series of three confirmatory factor analyses (CFA) to determine how well the teacher, parent, and student SSIS RFs fit the CASEL theoretical model. It should be emphasized that CFA is most appropriate for theory testing rather than for theory generation as is done in exploratory factor analysis (EFA). As per CFA procedures, the authors specified the entire factor structure by indicating which variables loaded and did not load on which CASEL factor and the degree to which the factors intercorrelated.

Once we established the factor structure for each of the SSIS SEL Edition RFs, we examined the reliability of their scores for the standardization subsample. Specifically, we documented the internal consistency, test-retest, and inter-rater reliability estimates for each RF. It was expected that these reliability estimates would meet or exceed common accepted criteria reliability (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education [AERA, APA, & NCME], 2014). We were primarily interested in whether or not our assigned items fit within the CASEL framework. Specifically, because the SSIS SEL Parent and Student forms were both designed to represent the five CASEL competency areas, we expected their CFAs would yield a five-factor solution representing self-awareness, self-management, social awareness, relationship skills, and responsible decision-making. Similarly, it was expected that the CFA for the SSIS SEL Teacher form would also yield these same five CASEL domains plus a sixth factor for academic competence.

Method

Participants

The SSIS SEL Edition used the teachers, parents, and students who participated in the original SSIS-RS standardization project in 2006–2007 (Gresham & Elliott, 2008). For that project, data were collected from a representative nationwide sample of 4,700 children and adolescents ages 3 through 18 years, who were assessed at 115 sites in 36 states. Demographic targets for the SSIS-RS norm sample were based on the *Current Population Survey, March 2006* (U.S. Census Bureau, 2006) and were applied to three norm groups: a preschool group (ages 3–5 years), and two

Table 1. Study Participant Characteristics.

	Teacher		Parent		Student	
Characteristic	Elementary	Secondary	Elementary	Secondary	Elementary	Secondary
n	146	75	126	114	139	85
Age						
M (years: months)	9:2	14:7	8:7	14:9	10:5	14:11
SD (months)	25	16.2	25	18.1	15.4	20.8
Sex						
Female	79	41	52	64	69	44
Male	67	34	74	50	70	41
Race/ethnicity						
African American	33	9	6	7	14	8
Hispanic	27	6	3	9	17	4
White	73	52	112	97	100	66
Other	13	8	5	1	8	7
Mother's education						
Grade 11 or less	16	4	4	2	11	4
Grade 12 or GED	52	23	34	29	46	16
1–3 yrs of college	50	26	43	41	46	35
4+ yrs of college	28	22	45	42	36	30
Region						
Northeast	11	9	45	30	9	21
North Central	28	19	41	38	54	36
South	66	17	34	26	57	6
West	41	30	6	20	19	22
Test interval (days)						
Range	30–72	0–89	32–90	29–88	56–89	26–86
M	53.4	29.2	60.2	63.1	68	60.7

Note. GED = General Educational Development.

school-age groups (ages 6–12 and 13–18 years). Each age group sample was designed to have equal numbers of males and females and to match the U.S. population with regard to race/ethnicity, socioeconomic status, and geographic region. Table 1 provides a detailed quantitative summary of our participants' characteristics.

For the reliability and validity studies, a subset of the overall standardization sample was selected. Teachers randomly selected and rated two students in their respective classrooms. The sample consisted of more than 200 ratings for teachers, parents, and students. Elementary teachers (n = 146; 35.5%) completed the teacher ratings. Parent ratings were completed by at least one parent, most often mothers, of 126 elementary students (30.6%) and 114 secondary students (41.6%). Elementary students (n = 139; 33.8%) and secondary students (n = 85; 31.0%) completed the self-report ratings. These are presented in Table 1, and the sample represented all major demographic categories.

A series of five chi-square analyses were performed to examine whether participants in our reliability and validity study subsample were significantly different from the participants in the overall norm sample of the SSIS. Results indicated that the sample for this study was representative of the national standardization sample used to develop the norms for the SSIS SEL Edition RFs. Specifically, there were no significant differences in terms of gender, age, race/ethnicity, geographic region, or socioeconomic status between the subsample and the national standardization sample ($p > .05$).

Procedures

Data collection. Site coordinators were recruited via the publisher's website and by contacting SSIS users and participants in other test development projects conducted by the publisher. The site coordinators responsibilities included working with various organizations to identify individuals to assess, to recruiting raters, serving as liaisons between classroom teachers and project staff, and managing the flow of data collection materials at their location. All site coordinators were required to meet the qualifications for using and interpreting the SSIS. Qualified individuals included school and clinical psychologists, special educators, school counselors, educational specialists, and other professionals such

as professors in graduate training programs or directors of research programs.

Quality control steps. To ensure success of the national standardization sampling, several quality control procedures were implemented before, during, and after data collection. Communication was maintained on a regular basis between the publisher and all site coordinators. Returned forms went through a series of check-in procedures. All completed forms were first checked against case assignment records to ensure the demographic information was completed correctly. Next, all case data were entered and verified. Forms with more than 10% of unscorable items (unmarked or multiply marked) were excluded from all analyses. The number of unscorable items was verified again on the keyed data files as part of the preparations for the final master file used for all subsequent analyses.

Measure

Description. The SSIS SEL Edition RFs consist of three multi-informant behavior rating forms (teacher, parent, and student) designed to assess each of the five competencies identified in the CASEL Five framework: Self-Awareness, Self-Management, Social Awareness, Relationship Skills, and Responsible Decision-Making (see previous descriptions and definitions). The SSIS SEL RS retained all social skills items plus five problem behavior items (reverse scored) from the SSIS-RS to reflect the CASEL framework for teaching SEL skills. CASEL's goal is to establish a unifying preschool through high school framework based on a coordinated set of evidence-based practices for improving social–emotional learning and academic performance (Weissberg et al., 2015). The CASEL conceptualization of SEL is widely accepted as a comprehensive and evidence-based approach to conceptualizing children and youths' social–emotional learning (CASEL, 2012, 2015; Durlak et al., 2011; Oberle, Domitrovich, Meyers, & Weissberg, 2016; Weissberg et al., 2015).

Adaptation method. The SSIS SEL Edition RFs are a reconfiguration of the SSIS-RS items to fit within the five CASEL SEL competency domains. Specifically, the SSIS SEL RFs are based on the consensus review of items by four SEL experts (the two authors of the rating scales, a consulting psychometrician to the publisher, and an outside consultant). The two authors have more than 65 years of collective experience in the development and use of social skills assessment and intervention tools. The psychometrician has more than 30 years of experience in the test development business, especially in the development of social behavior assessment tools. The outside consultant is a clinical psychologist with more than 38 years of experience in the field and a leadership role in the development and advancement of CASEL.

Each of the four SEL assessment experts reads the CASEL definitions of SEL competencies (i.e., self-awareness, self-management, social awareness, relationship skills, and responsible decision-making) as stated in the *Effective Social and Emotional Learning Programs: Preschool and Elementary School Edition* (CASEL, 2013). Using these SEL competency definitions, the four assessment experts independently read each of the 51 SSIS-RS items and assigned them to one of the five competency domains. This resulted in an initial agreement percentage of 74.5% (38 of 51 items) across all four persons. The two authors then reviewed each of the 13 items that did not have perfect agreement; three of the assessment experts agreed on the assignment of nine items, while for the remaining four items, only two assessment experts agreed on their assignment to an SEL competency domain. The competency domain assignment for the nine items that enjoyed majority agreement was accepted. With the remaining four items, the authors and psychometrician discussed their rationales for the item assignment until reaching an agreement of the best conceptual fitting CASEL domain.

For the teacher RF, we used the five CASEL domain definitions to assign social behavior items and used three additional domains to assign academically related items. Specifically, we used the following definitions to assign academic items: (a) Motivation to Learn is a state of arousal and actions directed toward learning and completing classroom tasks and activities; (b) Reading Skills are processes involving skills and development of subskills in interrelated domains such as meaning of words and phrases in context, understanding text, analyzing text, and evaluating and extending text; and (c) Mathematics Skills involve skills and development of subskills in interrelated domains including mathematical processes, number operations and relationships, geometry, measurement, statistics and probability, and algebraic relationships. The assignment of items for the SSIS SEL RF was identical to that of the SSIS-RS teacher version.

Scores on the five CASEL aligned SEL subscales are expressed as standard scores ($M = 100$, $SD = 15$). The teacher form consists of 51 SEL items distributed across the five CASEL domains and seven items that measure students' academic competence. The parent form consists of 51 SEL items, and the student form consists of 46 items (all social skills, no problem behavior items). Each of the SEL items is rated on a 4-point Likert-type scale measuring the frequency of the SEL skill (0 = *Never*, 1 = *Sometime*, 2 = *Often*, 3 = *Always*).

Data Analysis

CFAs. The scores from the SSIS SEL Teacher, Parent, and Student RFs are based on the national standardization sample of the SSIS-RS. These scores met the assumption of multivariate normality (for ungrouped data), did not have

Table 2. Internal Consistency Estimates for SSIS SEL Teacher Form.

	Ages 3–5 Years			Ages 5–12 Years			Ages 13–18 Years		
Scale	Female	Male	Combined	Female	Male	Combined	Female	Male	Combined
Self-Awareness	0.70	0.81	0.77	0.78	0.77	0.78	0.83	0.86	0.84
Self-Management	0.92	0.94	0.93	0.90	0.91	0.91	0.92	0.90	0.91
Social Awareness	0.87	0.88	0.89	0.88	0.93	0.91	0.93	0.89	0.91
Relationship Skills	0.89	0.92	0.91	0.88	0.91	0.90	0.93	0.91	0.92
Responsible Decision-Making	0.81	0.87	0.85	0.77	0.82	0.80	0.79	0.83	0.80
Academic Competence	—	—	—	0.97	0.97	0.97	0.97	0.94	0.96
SEL Composite Scale	0.96	0.97	0.97	0.96	0.96	0.96	0.97	0.97	0.97
Core Skills	0.89	0.91	0.91	0.90	0.91	0.91	0.90	0.91	0.90

Note. All scores reported in α. Dash in cell indicates scale not reported for age group. SSIS = Social Skills Improvement System; SEL = Social–Emotional Learning.

any outlying values, the data were homoscedastic and linear (determined by visual inspection of graphed data), and had no missing data based on the comprehensive data collection procedures.

Three CFAs were conducted using the AMOS version 22.0 subroutine in SPSS. In each of these analyses, we were exclusively interested in how well the teacher, parent, and student SSIS RFs fit the CASEL theoretical model. As per CFA procedures, the entire factor structure was specified by indicating which variables loaded and did not load on which CASEL factor and the degree to which the factors intercorrelated. To assess model fit, we used three fit indices: chi square, root mean square error of approximation (RMSEA), and comparative fit indices (CFI).

The five-factor Parent model showed a $\chi^2 = 14.274$ ($p < .0001$), the six-factor Teacher model showed a $\chi^2 = 11.225$ ($p < .0001$), and the five-factor Student model showed a $\chi^2 = 3.229$ ($p < .0001$). Although, in CFA, one desires a nonsignificant chi square, this statistic is unreasonable because it tests the hypothesis that the model fits perfectly in the population, which is highly unlikely. Also, the chi-square statistic is highly dependent on sample size, with larger samples almost always producing significant chi-square statistics (Bandalos, 2018; Hu & Bentler, 1998). For this reason, we did not use the chi-square statistic to assess model fit. Instead, we used the RMSEA and CFIs statistics to assess model fit because we were interested in evaluating how well our item assignment fit within the CASEL theoretical model. The five-factor Parent model produced a RMSEA statistic of .06 (90% confidence interval [CI] = [.061, .063]) and the five-factor Student model produced a RMSEA statistic of .06 (90% CI = [.052, .056]), demonstrating adequate model fit (Bandalos, 2018). The CFI values were .79 and .83 for the Parent and Student RFs, respectively. The six-factor Teacher model showed a RMSEA statistical value of .08 (90% CI = [.079, .82]) and the CFI value was .75, which was lower than the Parent and Student models, thereby demonstrating mediocre model fit. The results of these analyses are summarized in Table 2, and Figures 1 (Parent Form), 2 (Student Form), and 3 (Teacher Form).

Reliability estimates. Three methods were used to estimate the reliability of SSIS SEL RF's scores. Internal consistency estimates were computed using coefficient alphas for the teacher, parent, and student RFs.

For each SSIS SEL Edition RF, a sample of individuals was rated twice to establish an estimate of test-retest reliability. These samples were drawn from individuals who participated in the original SSIS standardization project, represented the full range of demographic characteristics (i.e., gender, race/ethnicity, parent education, and geographic region), and were consistent with the U.S. Census figures in these categories. The mean retest interval was 43 days for the Teacher form, 61 days for the Parent form, and 66 days for the Student form. The correlation of the scores across occasions for a sample of individuals rated twice, each time by the same rater, was used to estimate test-retest reliability coefficients. An adjusted correlation coefficient also is reported and provides a better estimate of the correlation that would be obtained on a fully representative sample by correcting for differences in score variance between the test-retest sample and the population (Allen & Yen, 2002). The adjusted coefficient is lower than the unadjusted coefficient when the score variance exceeds the variance in the norm group and is higher than the unadjusted coefficient when the variance in the sample is less than the variance in the norm group.

Finally, interrater reliability indicates the consistency of scores across two raters rating the same individual during a narrow time frame. Interrater reliability provides evidence of the level of consistency of scores that can be expected between pairs of raters (e.g., two teachers; two parents) who have a similar relationship to the individual they are rating. Two interrater reliability studies were conducted, a Teacher study and a Parent study.

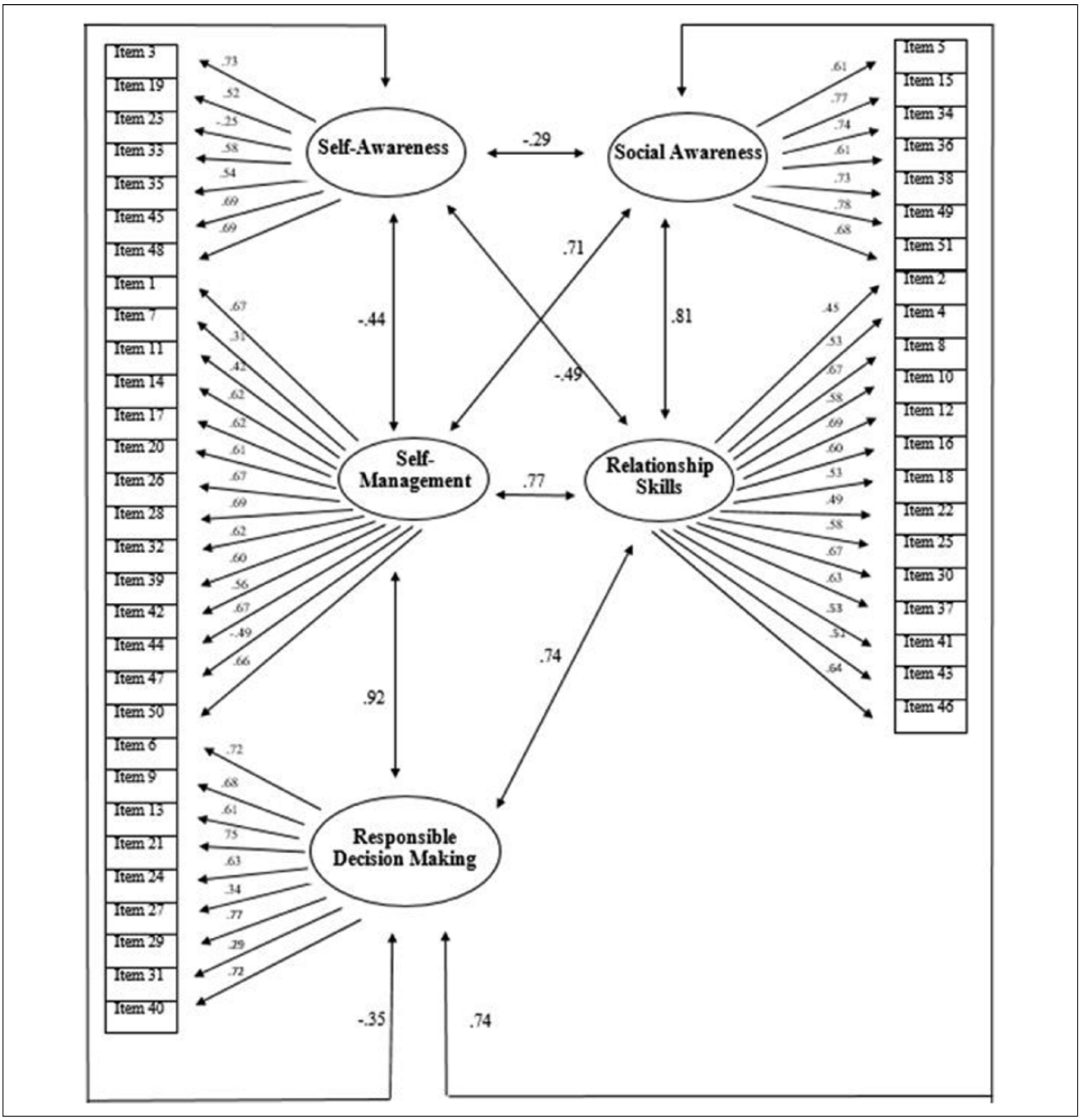

Figure 1. Five-factor model for SSIS SEL RF Parent.
Note. SSIS = Social Skills Improvement System; SEL RF = Social–Emotional Learning Rating Form.

Results

Evidence for Internal Structure for SSIS SEL RFs

The scores from the SSIS SEL Teacher, Parent, and Student RFs are based on the national standardization sample of the SSIS-RS. These scores met the assumption of multivariate normality (for ungrouped data), did not have any outlying values, the data were homoscedastic and linear (determined by visual inspection of graphed data), and had no missing data based on the comprehensive data collection procedures.

Three CFAs were conducted using the AMOS version 22.0 subroutine in SPSS. In each of these analyses, we were exclusively interested in how well the teacher, parent, and student SSIS RFs fit the CASEL theoretical model. As per

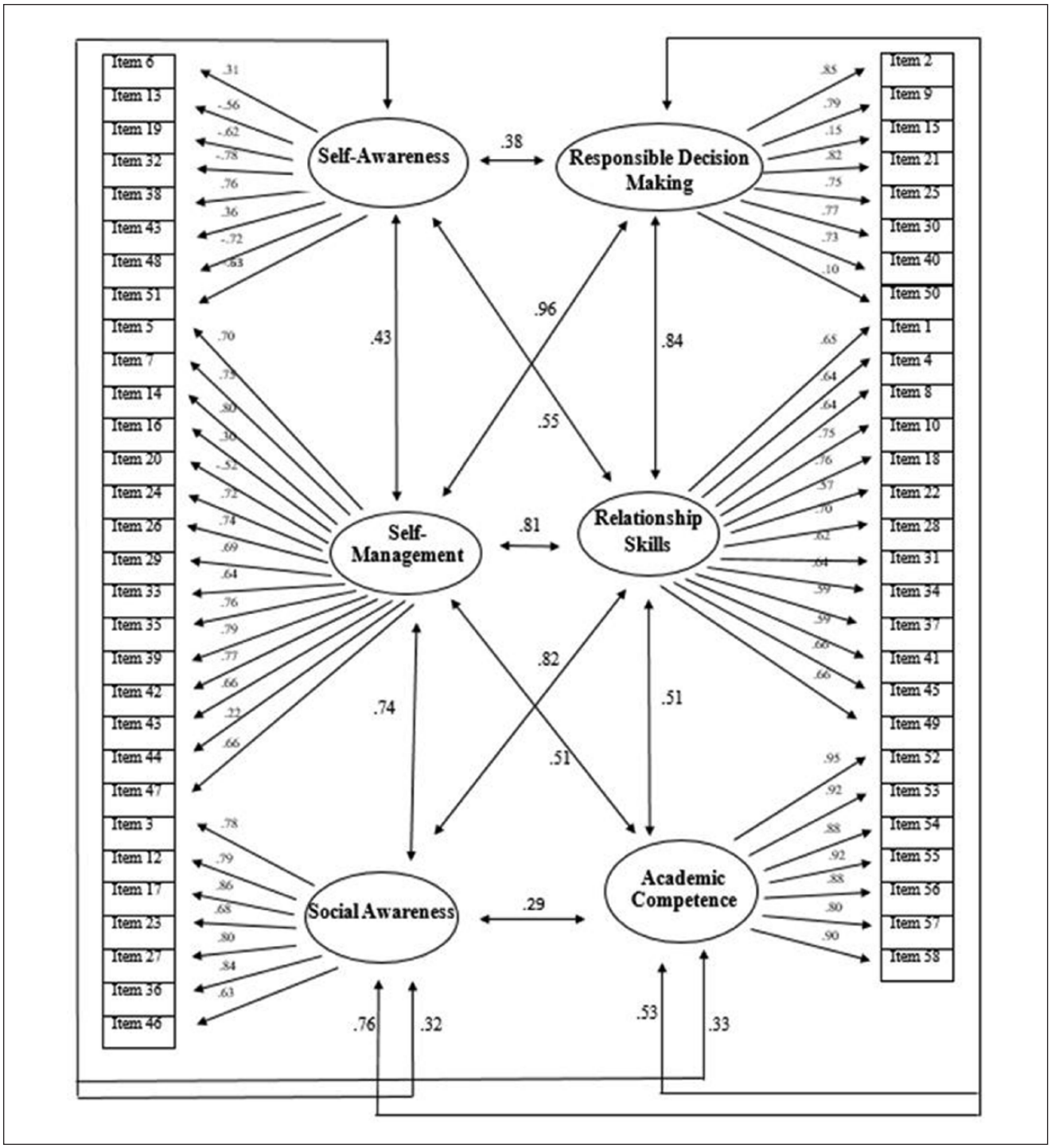

Figure 2. Six-factor model for SSIS SEL RF Teacher.
Note. SSIS = Social Skills Improvement System; SEL RF = Social–Emotional Learning Rating Form.

CFA procedures, the entire factor structure was specified by indicating which variables loaded and did not load on which CASEL factor and the degree to which the factors intercorrelated. To assess model fit, we report three fit statistics (chi square, RMSEA, and CFIs).

The five-factor Parent model showed a $\chi^2 = 14.274$ ($p < .0001$), the six-factor Teacher model showed a $\chi^2 = 11.225$ ($p < .0001$), and the five-factor Student model showed a $\chi^2 = 3.229$ ($p < .0001$). Although, in CFA, one desires a nonsignificant chi square, this statistic is unreasonable because

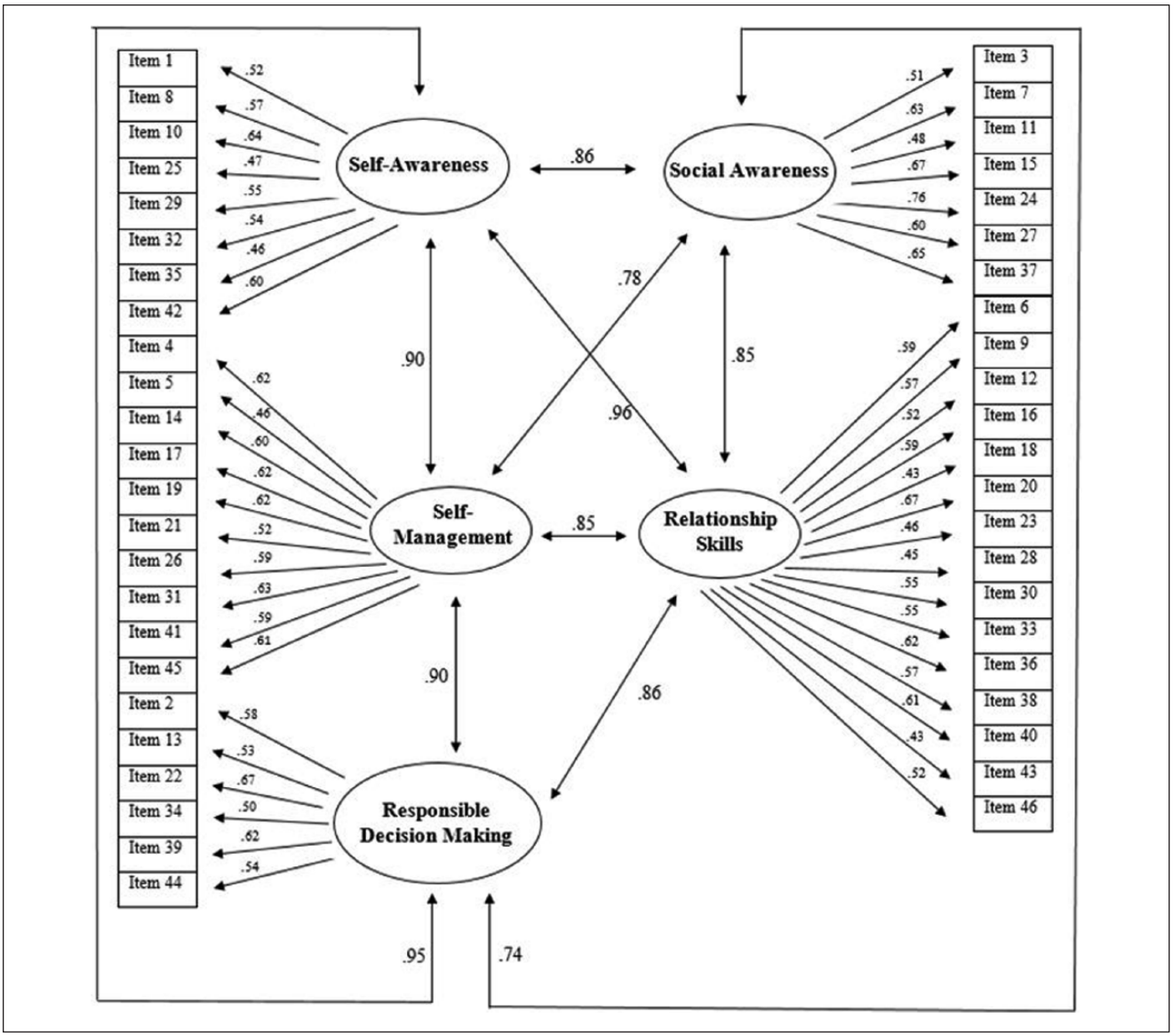

Figure 3. Five-factor model for SSIS SEL RF Student.
Note. SSIS = Social Skills Improvement System; SEL RF = Social–Emotional Learning Rating Form.

it tests the hypothesis that the model fits perfectly in the population, which is highly unlikely. Also, the chi-square statistic is highly dependent on sample size, with larger samples almost always producing significant chi-square statistics (Bandalos, 2018; Hu & Bentler, 1998). For this reason, we did not use the chi-square statistic to assess model fit. Instead, we used the RMSEA and CFIs statistics to assess model fit because we were interested in evaluating how well our item assignment fit within the CASEL theoretical model. The five-factor Parent model produced a RMSEA statistic of .06 (90% CI = [.061, .063]) and the five-factor Student model produced a RMSEA statistic of .06 (90% CI = [.052, .056]), demonstrating adequate model fit (Bandalos, 2018). The CFI values were .79 and .83 for the Parent and Student RFs, respectively. The six-factor Teacher model showed a RMSEA statistical value of .08 (90% CI = [.079, 0.82]), and the CFI value was .75, which was lower than the Parent and Student models, thereby demonstrating mediocre model fit. The results of these analyses are summarized in Table 2, and Figures 1 (Parent Form), 2 (Student Form), and 3 (Teacher Form).

Evidence for Internal Consistency Reliability

Teacher RF. Coefficient alphas are presented for male and females for the two age groups (5–12 years and 13–18 years) for the Teacher RF in Table 2. For each of the five CASEL subdomains, coefficient alphas for males and females ages 5

Table 3. Internal Consistency Estimates for SSIS SEL Parent Form: Normative Sample.

	Ages 3–5 Years			Ages 5–12 Years			Ages 13–18 Years		
Scale	Female	Male	Combined	Female	Male	Combined	Female	Male	Combined
Self-Awareness	0.74	0.76	0.75	0.76	0.75	0.75	0.78	0.82	0.80
Self-Management	0.86	0.89	0.88	0.87	0.87	0.87	0.87	0.88	0.88
Social Awareness	0.85	0.89	0.88	0.85	0.86	0.86	0.87	0.87	0.87
Relationship Skills	0.87	0.90	0.89	0.86	0.86	0.86	0.89	0.90	0.90
Responsible Decision-Making	0.83	0.88	0.86	0.81	0.83	0.82	0.8	0.84	0.83
SEL Composite Scale	0.95	0.96	0.96	0.95	0.95	0.95	0.95	0.96	0.96
Core Skills	0.83	0.88	0.86	0.83	0.83	0.83	0.84	0.88	0.86

Note. All scores reported in α. SSIS = Social Skills Improvement System; SEL = Social–Emotional Learning.

to 12 years ranged from 0.77 to 0.97, with a median value of 0.87. For 13 to 18 year olds, coefficient alphas for males and females ranged from 0.83 to 0.97, with a median value of 0.90. Coefficient alphas for the Academic Competence scale for males and females 5 to 12 years old was 0.97, and for the 13 to 18 age group was 0.97 for males and 0.94 for females. Coefficient alphas for males and females ages 5 to 12 years old for the SEL Composite Scale was 0.96, and for males and females ages 13 to 18 years was 0.97. Finally, on the Core Skills scale, coefficient alphas for males and females ages 5 to 12 years were 0.90 and 0.91, respectively, and for males and females ages 13 to 18 years were 0.90 and 0.91, respectively.

Parent RF. Coefficient alphas for males and females on the Parent RF subscales for 5 to 12 year olds ranged from 0.75 to 0.95, with a median value of 0.85. These values are presented in Table 3. Coefficient alphas for 5- to 12-year-old males and females using the Student RF ranged from 0.73 to 0.95, with a median value of 0.84. On the SEL Composite Scale, coefficient alphas for males and females ages 5 to 12 years were 0.95 and for males and females ages 13 to 18 years were 0.95. On the Core Skills Scale, coefficient alphas for males and females ages 5 to 12 years were 0.83 and for males and females ages 13 to 18 years were 0.84 and 0.88, respectively.

Student RF. Coefficient alphas using the Student RF subscales for males and females ages 5 to 12 years ranged from 0.73 to 0.95, with a median value of 0.84, and for males and females ages 13 to 18 years ranged from 0.83 to 0.95, with a median value of 0.89. These values are presented in Table 4. On the SEL Composite Scale, coefficient alphas for males and females ages 5 to 12 years were 0.94 and 0.95, respectively, and for males and females ages 13 to 18 years were 0.95. For the Core Skills, coefficient alphas for males and females for ages 5 to 12 years were 0.79 and for males and females ages 13 to 18 years were 0.87 and 0.83, respectively.

Evidence for Test-Retest Reliability

Teacher RF. For the SSIS SEL Edition Teacher RF, 144 individuals were rated twice by the same teacher. Table 5 presents the Teacher Form test-retest score means, standard deviations, and reliability coefficients for each administration. Corrected reliability coefficients were generally in the low .80s. Mean scores between administrations were very similar, with most effect sizes being under .10, indicating very stable ratings across the testing interval.

Parent RF. For the Parent Form, 115 individuals were rated twice by the same parent. Table 6 presents the Parent Form test-retest standard score means, standard deviations, and reliability coefficients for each administration. Corrected reliability coefficients were generally in the upper .70s and low .80s. Mean scores between administrations were very similar, with most effect sizes being under .10, indicating very stable ratings across the testing interval.

Student RF. For the Student Form, 127 individuals between the ages of 8 and 18 years rated themselves twice. Table 7 presents the Student Form test-retest standard score means, standard deviations, and reliability coefficients for each administration. Corrected reliability coefficients ranged from the mid .70s to the .80s. Mean scores between administrations were very similar, with most effect sizes being around .10, indicating very stable ratings across the testing interval.

Evidence for Interrater Reliability

Teacher RF. In the SSIS SEL Edition Teacher Form study, each individual was rated by pairs of raters (e.g., two teachers). Although all teachers had some interaction with the individual being rated, the level of interaction tended to be different. For example, among elementary school students, one rater may be the individual's teacher for most of the day, and the other rater may be the individual's teacher for one class period (e.g., a reading specialist or music teacher).

Table 4. Internal Consistency Estimates for SSIS SEL Student Form: Normative Sample.

	Ages 8–12 Years			Ages 13–18 Years		
Scale	Female	Male	Combined	Female	Male	Combined
Self-Awareness	0.73	0.78	0.77	0.84	0.82	0.83
Self-Management	0.79	0.83	0.81	0.85	0.84	0.85
Social Awareness	0.79	0.82	0.82	0.75	0.8	0.78
Relationship Skills	0.85	0.85	0.85	0.88	0.87	0.88
Responsible Decision-Making	0.74	0.70	0.72	0.78	0.75	0.76
SEL Composite Scale	0.94	0.95	0.94	0.95	0.95	0.95
Core Skills	0.79	0.79	0.79	0.87	0.83	0.85

Note. All scores reported in α. SSIS = Social Skills Improvement System; SEL = Social–Emotional Learning.

Table 5. Test-Retest Reliability for the Teacher Form.

Form and Scale	First Testing *M* (*SD*)	Second Testing *M* (*SD*)	*r*	Corrected *r*
Self-Awareness	100.2 (15.3)	100.0 (16.2)	.82	.81
Self-Management	99.6 (17.1)	99.9 (17.1)	.86	.82
Social Awareness	100.3 (16.5)	98.4 (15.5)	.78	.73
Relationship Skills	100.5 (15.9)	99.8 (16.2)	.83	.81
Responsible Decision-Making	102.1 (15.0)	100.7 (15.8)	.78	.78
Academic Competence	101.2 (15.9)	100.9 (15.8)	.92	.91
SEL Composite Scale	100.7 (16.0)	99.7 (16.0)	.84	.82
Core Skills	100.9 (16.0)	100.6 (16.9)	.84	.82

Note. The mean retest interval for the Teacher form was 43 days. SEL = Social–Emotional Learning.

Table 6. Test-Retest Reliability for the Parent Form.

Form and Scale	First Testing *M* (*SD*)	Second Testing *M* (*SD*)	*r*	Corrected *r*
Self-Awareness	100.1 (16.2)	101.8 (15.5)	.81	.78
Self-Management	98.5 (14.0)	99.6 (15.5)	.83	.85
Relationship Skills	97.4 (16.8)	97.6 (17.5)	.86	.82
Responsible Decision-Making	97.9 (15.0)	99.0 (14.8)	.79	.79
SEL Composite Scale	98.0 (16.6)	99.4 (16.3)	.87	.84
Core Skills	98.7 (15.0)	98.6 (16.0)	.84	.83

Note. The mean retest interval for the Parent form was 61 days. SEL = Social–Emotional Learning.

Table 7. Test-Retest Reliability for the Student Form.

Form and Scale	First Testing *M* (*SD*)	Second Testing *M* (*SD*)	*r*	Corrected *r*
Self-Awareness	98.5 (14.9)	101.5 (14.4)	.73	.73
Self-Management	99.3 (14.3)	101.4 (15.0)	.71	.74
Relationship Skills	99.2 (13.9)	99.7 (14.3)	.68	.73
Responsible Decision-Making	99.8 (13.8)	101.2 (13.7)	.77	.81
SEL Composite Scale	99.2 (14.3)	100.8 (14.5)	.81	.83
Core Skills	99.4 (14.4)	101.4 (15.0)	.75	.77

Note. The mean retest interval for the Student form was 66 days. SEL = Social–Emotional Learning.

Table 8. Interrater Reliability for the Teacher Form.

Scale	First Rater *M* (*SD*)	Second Rater *M* (*SD*)	*r*	Corrected *r*
Self-Awareness	99.2 (17.2)	95.9 (16.5)	.53	.38
Self-Management	101.1 (15.9)	98.1 (13.6)	.67	.63
Social Awareness	99.5 (14.4)	95.7 (16.4)	.56	.59
Relationship Skills	99.5 (17.6)	99.0 (13.9)	.72	.61
Responsible Decision-Making	101.5 (15.8)	99.9 (13.9)	.62	.55
Academic Competence	100.2 (16.4)	96.7 (15.2)	.62	.55
SEL Composite Scale	100.2 (17.1)	96.7 (15.5)	.69	.60
Core Skills	100.8 (15.4)	97.4 (13.6)	.70	.68

Note. SEL = Social–Emotional Learning.

Table 9. Interrater Reliability for the Parent Form.

Scale	First Rater *M* (*SD*)	Second Rater *M* (*SD*)	*r*	Corrected *r*
Self-Awareness	101.2 (13.5)	99.1 (13.7)	.34	.47
Self-Management	99.5 (15.2)	101.1 (14.6)	.63	.62
Social Awareness	99.3 (14.4)	98.3 (14.6)	.53	.57
Relationship Skills	99.5 (15.0)	98.3 (14.3)	.60	.60
Responsible Decision-Making	99.8 (14.5)	98.2 (14.7)	.59	.62
SEL Composite Scale	99.9 (14.4)	98.5 (14.6)	.62	.65
Core Skills	98.7 (14.3)	98.3 (14.5)	.63	

Note. SEL = Social–Emotional Learning.

In some cases, a paraprofessional or other school staff member who had regular opportunities to observe the student several times a week in a classroom or other organized group setting (e.g., an afterschool program, special pull-out program, etc.) supplied the second rating.

For the Teacher Form, 54 students were rated by two teachers. The sample included students from each of the demographic categories of gender, race/ethnicity, parent's education, and geographic region. The mean interval between ratings was 63 days. Table 8 presents the Teacher Form interrater standard score means, standard deviations, and reliability coefficients. Corrected reliability coefficients ranged from the upper .30s to the .60s, with a median reliability coefficient of .53. Mean score differences between administrations were generally small, with most effect sizes being around .20, indicating relatively consistent views of behavioral performance across raters. Ratings provided by the second rater were slightly lower than the ratings provided by the first rater.

Parent RF. In the Parent Form study, each individual was rated either by both parents or a parent and a close relative/caregiver. The raters had regular opportunities to observe the individual's interactions with members of the family and/or community both at home and in other social settings. For the Parent Form, 110 students were rated by two caregivers. The mean interval between ratings was 58 days. Table 9 presents the Parent Form interrater standard score means, standard deviations, and reliability coefficients. Corrected reliability coefficients ranged from the upper .40s to the lower .70s with a mean reliability coefficient of .59. Mean score differences between administrations were generally small, with all effect sizes being .15 or less, indicating consistent views of behavioral performance across raters.

Discussion

This study provided a brief description of key aspects of the development of a new assessment of children's SEL skills, the SSIS SEL RFs, and an examination of fundamental psychometrics of these assessments' internal structure and reliability estimates. Specifically, the SSIS SEL Edition was designed to directly align with the SEL competencies advanced by CASEL. The CASEL SEL model has influenced the definition of SEL used by many state departments of education in the United States, as well as in numerous foreign countries (e.g., Australia, Canada, Denmark, England, New Zealand, and Singapore) and is prevalent in universal intervention programs in many U.S. schools. The

alignment of content between assessments and interventions contributes information that is more content relevant and representative. Thus, the utility of the assessments is likely enhanced, assuming the assessments yield reliable and valid scores.

Major Findings

Two major findings about the SSIS SEL Teacher, Parent, and Student RFs were deduced from our psychometric results. Collectively, this evidence provides users with initial validity and reliability information called for in the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) for norm-referenced assessments used to identify students' strengths and weaknesses and evaluate outcomes of interventions designed to improve these students' social behaviors.

First, we found that the statistical analyses with the SSIS SEL Edition indicated the RF for the Parents and Students nicely fit the expected structure theorized by CASEL's five SEL competency domains. The Teacher version fits less well with the CASEL model and is best characterized as mediocre. The Teacher version also clearly captured a sixth factor for the academic competence items only included for teachers to rate. Remember, our purpose was only to determine whether or not our assigned items fit within the CASEL framework, which is a legitimate and accepted use of CFA procedures (Bandalos, 2018; Hu & Bentler, 1998). Specifically, the CFA demonstrated that the social skills items from the original SSIS-RS Teacher, Parent, and Student versions could meaningfully be reorganized to represent the SEL subconstructs of self-awareness, self-management, social awareness, relationship skills, and responsible decision-making. This is an important finding for two reasons. It provides evidence for the multidimensional scoring interpretation framework used with the assessments, and it establishes the SSIS SEL RF as a theoretically and empirically derived measure of the CASEL five model of SEL competency. Other assessments such as the DESSA (LeBuffe, Shapiro, & Naglieri, 2009) have scales that also measure some of the skills theorized as critical aspects of the CASEL model of SEL competencies.

Second, we found a family of reliability estimates for the three SSIS SEL RFs' scales and subscales for male and female students all met accepted criteria (Bandalos, 2018). Specifically, we examined the internal consistency, test-retest, and interrater reliability of scores for students from the national standardization sample. A recap of the results indicates the SEL Composite scale consistently produced high coefficient alphas, indicating that measures of aggregate social–emotional skills and their subscales across the Teacher, Parent, and Student forms are comparably reliable. Similarly, the Academic Competence Scale on the SSIS SEL Teacher yielded a high coefficient alpha. For subsamples of students representative of the entire standardization sample, we also found that Teachers' and Parents' SEL ratings at the scale and subscale levels were very stable over periods of 4 to 6 weeks. Finally, when using pairs of teachers and sets of parents/caregivers familiar with their students, we found that the interrater reliabilities for both the Teacher and Parent forms was in the expected moderate range, indicating good overall agreement between two raters of children's SEL skills. These internal consistency and test-retest reliability estimates are very similar to those reported for the original SSIS-RS (Gresham & Elliott, 2008). The interrater results were directly compared with the SSIS-RS interrater results and found to be very similar and superior to what is typically reported for social behavior rating scales (De Los Reyes & Kazdin, 2005).

Study Limitations and Future Research

Although the analyses offer some meaningful insights into the technical aspects of the new SSIS SEL assessments, our CFAs for the Teacher, Parent, and Student RFs were each only tested against the influential CASEL Five Competence model and were based on an extant pool of items. Improved fit statistics and perhaps a stronger test of the CASEL model could be advanced with the addition of some new items directly influenced by the CASEL definitions. In addition, our sample sizes for these CFAs, although adequately sized when we aggregated all students for each form, did not allow for further age group breakdowns. Another potential limitation is that the data for the SSIS RFs collected a decade ago may underestimate the minority student population to some extent. These limitations are important to address with future research.

Additional areas for future research with the SSIS SEL Edition RFs include validity evidence for subgroups of children such as those identified with emotional behavior disorders and autism. Children with these types of difficulties are likely to be assessed and perhaps served with interventions for social–emotional difficulties. Finally, research is encouraged to better understand the relationship between SEL competencies and academic functioning. Much has been made of the correlational findings that SEL skills function as academic enablers (e.g., DiPerna et al., 2002, 2005), but more research needs to be conducted to establish the causal and longitudinal relationship between these constructs.

Conclusion

The goal for the development of the SSIS SEL Edition RFs is to continue the SSIS tradition of reliable and valid multirater assessment of children's social–emotional behavior while aligning these assessments with the CASEL SEL

competency framework. Thus, these new RFs were confirmed to fit the five theoretical SEL domains—self-awareness, self-management, social awareness, relationship skills, and responsible decision-making—advanced by CASEL. The fit of the CASEL model was better for the Parent and Student RFs than that found with the Teacher RF.

Educators and psychologists interested in measuring and improving SEL skills of children ages 3 to 18 now have well-aligned, online assessments with the CASEL SEL skill definitions that offer norm-referenced scores for males and females individually or as a group. The findings in this study also provided substantial evidence for the reliability of the resulting scores from the SSIS SEL RFs when completed by teachers, parents, and students alike. In summary, with the rapidly expanding SEL research literature and growth of school-based SEL interventions, assessments such as the SSIS SEL RFs can contribute reliable and valid information that educators need to make sound decisions about children's social–emotional learning skills.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

Allen, M., & Yen, W. (2002). *Introduction to measurement theory*. Long Grove, IL: Waveland Press.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Author.

Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. New York, NY: Guilford Press.

Collaborative for Academic, Social, and Emotional Learning (2012). *2013 CASEL Guide: Effective social and emotional learning programs (Preschool and elementary editions)*. Chicago, IL: Author.

Collaborative for Academic, Social, and Emotional Learning. (2013). *Effective social and emotional learning programs: Preschool and elementary school edition*. Chicago, IL: Author.

Collaborative for Academic, Social, and Emotional Learning. (2015). *Effective social and emotional learning programs: Middle school and high school*. Chicago, IL: Author.

Conduct Problems Prevention Research Group. (1999). Initial impact of the Fast Track prevention trial for conduct problems: I. The high-risk sample. *Journal of Consulting and Clinical Psychology*, *67*, 631–647.

Conduct Problems Prevention Research Group. (2002). Evaluation of the first 3 years of the Fast Track prevention trial with children at high risk for adolescent conduct problems. *Journal of Abnormal Child Psychology*, *30*, 19–35.

Crowe, L. M., Beauchamp, M. H., Catroppa, C., & Anderson, V. (2011). Social function assessment tools for children and adolescents: A systematic review from 1988 to 2010. *Clinical Psychology Review*, *31*, 767–785.

De Los Reyes, A., & Kazdin, A. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*, 483–509.

DiPerna, J., Volpe, R., & Elliott, S. N. (2002). A model of academic enablers and elementary reading/language arts achievement. *School Psychology Review*, *31*, 298–312.

DiPerna, J., Volpe, R., & Elliott, S. N. (2005). A model of academic enablers and mathematics achievement in elementary grades. *Journal of School Psychology*, *43*, 379–392.

Durlak, J., Weissberg, R., Dymnicki, A., Taylor, R., & Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405–432.

Elliott, S. N., Frey, J., & Davies, M. (2015). Systems for assessing and improving students' social skills to achieve academic competence. In J. Durlak, C. Domitrovich, R. Weissberg & T. Gullotta (Eds.), *Handbook for social and emotional learning* (pp. 301–319). New York, NY: Guilford Press.

Epstein, M. H. (1999). The development and validation of a scale to assess the emotional and behavioral strengths of children and adolescents. *Remedial and Special Education*, *25*, 258–263.

Epstein, M. H. (2004). *Behavioral and Emotional Rating Scale–2nd Edition: A strength based approach to assessment*. Austin, TX: PRO-ED.

Greenberg, M., Weissberg, R., O'Brien, M., Zins, J., Fredericks, L., Resnik, H., & Elias, M. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, *58*, 466–474.

Gresham, F. M. (2018). *Effective interventions for social-emotional learning*. New York, NY: Guilford Press.

Gresham, F. M., Cook, C., Crews, S. D., & Kern, L. (2004). Social skills training for children and youth with emotional and behavioral disorders: Validity considerations and future directions. *Behavioral Disorders*, *30*, 19–33.

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating scales*. Bloomington, MN: Pearson Assessments.

Gresham, F. M., & Elliott, S. N. (2008). *Social Skills Improvement System Rating Scales*. Bloomington, MN: Pearson Assessments.

Gresham, F. M., & Elliott, S. N. (2017). *Social skills improvement system social–emotional learning edition rating forms*. Bloomington, MN: Pearson Assessments.

Gresham, F. M., Elliott, S. N., Vance, M. J., & Cook, C. R. (2011). Comparability of the Social Skills Rating System to the Social Skills Improvement System and psychometric comparisons across elementary and secondary age levels. *School Psychology Quarterly, 26*, 27–44.

Gresham, F. M., MacMillan, D., Bocian, K., Ward, S., & Forness, S. (1998). Comorbidity of hyperactivity-impulsivity-inattention and conduct problems: Risk factors in social, affective, and academic domains. *Journal of Abnormal Child Psychology*, *26*, 393–406.

Gresham, F. M., & Reschly, D. J. (1987). Dimensions of social competence: Method factors in the assessment of adaptive

behavior, social skills, and peer acceptance. *Journal of School Psychology*, *25*, 367–381.

Harniss, M. K., Epstein, M. H., Ryser, G., & Pearson, N. (1999). The Behavioral and Emotional Rating Scale: Convergent validity. *Journal of Psychoeducational Assessment*, *17*, 4–14.

Hu, L., & Bentler, P. (1998). Fit indices in covariance structure modeling: Sensitivity to overparameterized model misspecification. *Psychological Methods*, *3*, 424–453.

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lentrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people. *Educational and Psychological Measurement*, *71*, 617–637.

Humphrey, N., Lendrum, A., Wigelsworth, M., & Greenberg, M. (2016). Editorial introduction. *Cambridge Journal of Education*, *46*, 271–275.

LeBuffe, P. A., Shapiro, V., & Naglieri, J. (2014). *The Devereux Student Strengths Assessment (DESS). Assessment and technical manual and user's guide*. Charlotte, NC: Apperson. (Original work published 2009)

Maag, J. (2005). Social skills training for youth with emotional and behavioral disorders and learning disabilities: Problems, conclusions, and suggestions. *Exceptionality*, *13*, 155–172.

Merrell, K. (2011). *Social and emotional assets scale (SEARS)*. Lutz, FL: Psychological Assessment Resources.

MTA Cooperative Group. (1999). A 14-month randomized clinical trial of treatment strategies for attention–deficit/hyperactivity disorder (ADHD). *Archives of General Psychiatry*, *56*, 1073–1086.

Naglieri, J. A., Lebuffe, P., & Shapiro, V. B. (2011). *Devereux Student Strengths Assessment Mini-DESSA Assessment, Technical Manual, and User's Guide*. Villanova, PA: Apperson.

Nese, R., Doerner, E., Romer, N., Kaye, N., Merrell, K., & Tom, K. (2012). Social and Emotional Assets and Resilience Scales: Development of a strength-based short-form behavior rating scale system. *Journal for Educational Research Online*, *4*, 124–139.

Nickerson, A., & Fishman, C. (2009). Convergent and divergent validity of the Devereux Student Strengths Assessment. *School Psychology Quarterly*, *24*, 48–59.

Oberle, E., Domitrovich, C., Meyers, D., & Weissberg, R. (2016). Establishing systemic and social emotional learning approaches in schools: A framework for schoolwide implementation. *Cambridge Journal of Education*, *46*, 277–298.

Organisation for Economic Co-operation and Development. (2015). *Skills for social progress: The power of social and emotional skills*. Paris, France: Skills Studies, Organisation for Economic Co-operation and Development.

Shapiro, V., Accomazzo, S., & Robitalle, J. (2017). In the same ballpark or a whole new ball game? Staff as raters of youth behavior. *Journal of Child and Family Studies*, *26*, 1051–1055.

Shapiro, V., Kim, B., Robitaille, J., & LeBuffe, P. (2017). Protective factor screening for prevention practice: Sensitivity and specificity of the DESSA-Mini. *School Psychology Quarterly*, *32*, 449–464.

Taylor, R. D., Oberle, E., Durlak, J., & Weissberg, R. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, *88*, 1156–1171.

U.S. Census Bureau. (2006). *Current population survey*. Washington, DC: Author.

Walker, H. M. & Gresham, F. M. (Eds.). (2014). *Handbook of evidence-based practices for emotional and behavioral disorders: Applications in schools*. New York, NY: Guilford Press.

Weissberg, R., & Cascarino, J. (2013). Academic learning + social-emotional learning = national priority. *Phi Delta Kappan*, *95*, 8–13.

Weissberg, R., Durlak, J., Domitrovich, C., & Gullotta, T. (2015). Social and emotional learning: Past, present, and future. In J. Durlak, C. Domitrovich, R. Weissberg & T. Gullotta (Eds.), *Handbook for social and emotional learning* (pp. 3–19). New York, NY: Guilford Press.

World Economic Forum. (2016, March). *New vision for education: Fostering social and emotional learning through technology*. Geneva, Switzerland: Author.