

Gresham, F.M., Elliott, S.N., Byrd, S., Wilson, E., & Cassidy, K. (2018). Cross-informant agreement of children's social emotional skills: An investigation of ratings by teachers, parents, and students from a nationally representative sample. *Psychology in the Schools*, 55 (2), 208-223. doi: 10.1002/pits.22101

**Cross-Informant Agreement of Children's Social Emotional Skills:  
An Investigation of Ratings by Teachers, Parents, and Students  
From a Nationally Representative Sample**

**IMPACT STATEMENT**

Users of multi-rater assessments like the SSIS SEL Edition Rating Forms can expect to find some disagreement across raters because situations and environments influence social behavior. Multiple raters can be expected to rate the frequency of some social emotional skills somewhat differently, however, the present study indicated much more agreement than disagreement when focusing on composite scores.

### **Abstract**

This study examined the agreement across informant pairs of teachers, parents, and students regarding the students' social emotional learning competencies. Two student subsamples representative of the SSIS SEL Edition Rating Forms national standardization sample were examined; first 168 students (3<sup>rd</sup> to 12<sup>th</sup> grades) with ratings by three informants (a teacher, a parent, and the student him/herself) and a second group of 164 students who had ratings by two raters in a similar role – two parents or two teachers. To assess inter-rater agreements, two methods were employed; calculation of  $q$  correlations among pairs of raters and effect size indices to capture the extent rater pairs differed in their assessments of social-emotional skills. The empirical results indicated that pairs of different types of informants exhibited greater than chance levels of agreements as indexed by significant interrater correlations, teacher-parent informants showed higher correlations than teacher-student or parent-student pairs across all SEL competency domains assessed, and pairs of similar informants exhibited significantly higher correlations than pairs of dissimilar informants. Study limitations are identified and future research needs outlined.

**Cross-Informant Agreement of Children's Social Emotional Skills:  
An Investigation of Ratings by Teachers, Parents, and Students  
From a Nationally Representative Sample**

Over the past three decades, the importance of children's social competence has become increasingly clear across a variety of settings and outcomes. Whether this competence is referred to as social skills, personal adjustment, social functioning, social behavior, or social-emotional learning, the findings have been consistent: Children and youth with stronger skills and higher functioning levels tend to have greater academic success, tend to be more socially adjusted, and have lower risk of serious psychopathology in adulthood (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2001; Elliott, Davies, & Frey, 2015). Social-emotional learning (SEL) is a process through which individuals acquire and apply the knowledge, attitudes, and skills necessary to manage emotions, set and achieve goals, feel and show empathy, establish and maintain positive relationships, and make responsible decisions (Collaborative on Academic Social Emotional Learning, 2017). This process is one that occurs both in and out of school, from early years through adulthood. As noted by the Collaborative on Academic Social Emotional Learning (CASEL, 2015), SEL is a core element of academic success, and well-designed and coordinated efforts aimed at increasing knowledge and skills can be highly effective.

Substantial correlational and longitudinal research evidence shows that children's social and emotional competencies are related to positive adjustment outcomes and inversely related to a number of problem behaviors (e.g., Weissburg, Durlak, Domitrovich, & Gullota, 2015). Researchers also have documented a relatively strong predictive relationship between children's

social-emotional behaviors and their long-term academic achievement (Capara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000; DiPerma & Elliott, 2002; Malecki & Elliott, 2002). Specifically, the concept of social-emotional skills as academic enablers evolved from research that documented the relationship between students' nonacademic behaviors (e.g., social skills, engagement behaviors, motivation) and their academic achievement (DiPerna & Elliott, 2002; Gresham & Elliott, 2008; Wentzel, 1993).

The meta-analysis by Durlak and colleagues (2011) of 213 studies, which included over 270,000 students, documented that SEL intervention programs often lead to significant positive outcomes in several important areas of adjustment. These outcomes included (a) improvements in academic performance, (b) improved SEL skills, (c) improved prosocial behaviors and attitudes, and (d) reductions in conduct problems and emotional distress (e.g., anxiety and depression). The magnitude of the effect sizes for these interventions ranged from 0.22 to 0.57, depending on the outcome measure. These effect sizes are comparable to other well-established psychosocial interventions.

A subsequent meta-analysis by Taylor, Oberle, Durlak, and Weissberg (2017) that examined follow-up effects for up to 18 years for 82 of these SEL interventions demonstrated SEL's enhancement of positive youth development. Specifically, participants in SEL interventions "fared significantly better than controls in social emotional skills, attitudes, and indicators of well-being. Benefits were similar regardless of students' race, socioeconomic background, or school location" (p. 1156).

Despite the recognition that SEL skills are important and numerous interventions exist that can effectively improve these skills, there is a need for better assessments of SEL skills. In particular, there is a need for assessments that are psychometrically sound and **can be used by**

multiple raters to identify children's and youth's SEL skill strengths and weaknesses (McKown, 2017). The need for multiple rater strategies is very relevant to sound assessments across settings of school, home, and community.

### **Assessment of SEL Skills: Advantages of Behavior Rating Scales**

Behavior rating scales are among the most frequently used measures of social behavior in school and clinical settings (Achenbach & Rescorla, 2001; Gresham & Elliott, 2008, 2017; Kamphaus & Reynolds, 2016). Behavior rating scales are considered indirect measures of behavior because they are not measuring behavior at a time and place of its actual occurrence. Instead, behavior ratings require the rater to retrospectively reflect and rate the occurrence of specific behaviors. Behavior rating scales have several advantages, including: (a) information is quantifiable and amenable to reliability and validity analyses, (b) broad range of behavior (e.g., social skills and problem behaviors) can be assessed, (c) multiple raters can be used to assess behavior from multiple perspectives (teachers, parents, students), and (d) normative data provide a standard for judging the severity of behavior by comparing an individual with representative samples of other individuals (Gresham & Elliott, 2008; McConaughy & Ritter, 2005).

**Behavior rating scales have a long history of demonstrating of reliability and validity evidence.** Some of the most common broadband behavior ratings scales are well established and have demonstrated technical adequacy in terms of reliability and validity. These measures include the Behavioral Assessment System for Children-3 (BASC-3; Kamphaus & Reynolds, 20016), Achenbach System of Empirically Based Assessment (ASEBA; Achenbach & Rescorla, 2001), and the Social Skills Improvement System Rating Scales (Gresham & Elliott, 2008). A review of this literature is beyond the scope of this article, but it is sufficient to conclude that these measures are psychometrically adequate and are widely used by

professionals in the field (Humphrey, Kalambouka, Wigelsworth, Lendrum, Deighton, & Wolpert, 2011).

### **Behavior Rating Scales and Informant Discrepancy Research**

Informant discrepancies have a significant impact on assessment, classification, and treatment of social behavioral difficulties. Reliance on different informants can lead to identifying different children in a given population as meeting criteria for a given disorder (De Los Reyes & Kazdin, 2005). For example, the prevalence rate of conduct and oppositional disorders in community samples range from 1.6% to 10.2% depending on whether parent or teacher ratings are used to classify these disorders or whether both are considered simultaneously (Offord, Boyle, Racine, Szatmari, Fleming, Sanford, et al., 1996). Prevalence of classification of disorders ranges widely in clinical samples as well. Using parent or teacher ratings, or combining information from multiple informants from both, the prevalence of conduct disorder ranges from 9.7% to 23% and emotional disorder (anxiety, depression) ranges from 10.3% to 36.2% (MacLeod, McNamee, Boyle, Offord, & Friderich, 1999).

Source and context variance refers to systematic effects specific to a certain source of information (i.e., teachers, parents, self) within a given social context. Traditionally, source effects have been assumed to reflect bias associated with characteristics of the rater. For example, teacher ratings reflect their observations and perceptions of a child based on a few months in school-related contexts, **whereas parent ratings are based on a much longer history of observation of their own children in home other social contexts.**

Another view of source and context effects is that they reflect real differences in children's behavior across settings/situations as perceived by informants familiar with the child in these different settings. For example, differences in ratings of inattention could result from a

child who frequently displays inattention in the classroom in the presence of the teacher, but rarely displays such behavior at home in the presence of the mother. Evidence from two separate lines of ADHD research supports this alternative view of source variance: systematic multitrait and multisource investigations of the construct validity of ADHD rating scales (see Valo & Tannock, 2010 for a comprehensive review). One line of this research indicates that ADHD symptoms are situationally specific and do not reflect real difference in the way that parents and teachers interpret ADHD symptoms. The second line of research indicates that parent reports of symptoms at home have limited utility in predicting teacher reports of children's functioning at school. Based on these two lines of research, the general consensus is that a diagnosis of ADHD and its subtypes based on information from a parent alone may be of questionable validity because information obtained directly from teachers provides diagnostically relevant information (Valo & Tannock, 2010).

Virtually all of the research on informant discrepancies has been conducted using behavior rating scales that measure problem behaviors such as ADHD, conduct disorders, anxiety disorders, or depression (De Los Reyes & Kazdin, 2005). One exception is the study by Gresham and colleagues who used the national standardization data of the Social Skills Improvement System Rating Scales (SSIS-RS) to compute cross-informant correlations across various social skills and problem behavior domains (Gresham, Elliott, Cook, Vance, & Kettler, 2010). For total social skills ratings, these correlations were .30, .21, and .21 for teacher-parent, parent-student, and teacher-student ratings, respectively. These values are similar to the cross-informant correlations reported by Gresham and Elliott (1990) who used the national standardization data from the Social Skills Rating System. It should be noted, however, that these correlations were not computed on identical item content across the three raters. In fact, the

SSIS-RS teacher-parent forms have 80% of the items in common, while the teacher-student forms share only 67% of the items in common. The issues of using common items across rating forms and controlling for construct irrelevant variance are discussed later in this article.

### **Purpose of the Present Study**

The purpose of the present study was to systematically explore patterns of agreement among teachers, parents, and students across several SEL competence domains. This investigation used the revision of the SSIS Social Emotional Learning Edition Rating Forms (Gresham & Elliott, 2017). The SSIS SEL Edition Rating Forms represents a multi-rater (teacher, parent, and student) series of rating forms to assess the frequency of SEL skills (see details regarding these assessments in the Instrumentation section). Three predictions based on previous cross-informant agreement research were specifically tested in this investigation:

1. **Pairs of informants** (teacher-parent, teacher-student, and parent-student) exhibit greater than chance levels of agreements as indexed by significant interrater Pearson  $r$  correlations. These correlations were expected to be greater than those reported by Achenbach et al.'s (1987) meta-analysis and Gresham et al.'s (2010) study using the Social Skills Improvement System Rating Scales (SSIS-RS).
2. Teacher-parent informants show higher Pearson  $r$  correlations than teacher-student and parent-student pairs across SEL domains.
3. Pairs of similar informants (teacher-teacher and parent-parent) show higher Pearson  $r$  correlations than pairs of dissimilar informants (teacher-parent, teacher-student, parent-student).

To empirically test these predictions, we used rating scale data from the SSIS national



standardization norming sample to provide evidence for a large and representative sample of students from across the United States. The details of our method follow.

## **Method**

### **Participants**

Two subsamples from the standardization sample of the SSIS-Rating Forms SEL Edition were used (Gresham & Elliott, 2017). The first sample of participants consisted of 168 students who had all three informants (teacher, parent, and student) complete ratings of their social behaviors on respective SEL Edition SSIS-Rating Forms. Table 1 depicts the demographic information for these 168 participants. The average age of the participants was 11.9 years with the majority being male (63%). A series of Chi square ( $\chi^2$ ) analyses were performed to examine whether these 168 participants were significantly different than the other 4,382 participants in the norm-referenced sample. Results indicated that the sample for this study was representative of the national standardization sample used to develop norms for the SSIS-Rating Forms SEL Edition. Specifically, there were no significant differences in terms of gender, grade or age between the subsample for this cross-informant study and the entire standardization sample.

The second subsample of participants consisted of 164 students who had similar raters (i.e., teacher-teacher and parent-parent) complete the SSIS SEL Edition Rating Forms that was necessary to assess the agreement between similar raters. Table 3 depicts the demographic information for these participants. The average age of these participants was 9 years and the majority of the participants were male and of White racial background. Similar to the first subsample, a series of  $\chi^2$  analyses indicated that the sample for this study was representative of the national standardization sample used to develop the SSIS-Rating Forms SEL Edition norms.

A power analysis was computed for Pearson correlation coefficients with a power of 0.80 and  $p < .01$  and indicated a sample size of  $N = 118$  was need. The size of both subsamples in the study clearly exceeded this number.

### **Instrumentation**

The SEL Edition of the SSIS-Rating Forms offers a multi-rater series of rating scales for teachers, parents, and students (Gresham & Elliott, 2017). All rating forms include items for each of the five SEL competency domains identified by the Collaborative for Academic Social and Emotional Learning (CASEL, 2015). These competencies are: Self-Awareness, Self-Management, Social Awareness, Relationship Skills, and Responsible Decision Making.

The SSIS SEL Rating Forms were designed based on the principle of alignment of assessment content (as operationalized with skill-focused items) with intervention content (skills taught) to improve construct representation. As such, minimizing or avoiding problems of construct underrepresentation or construct misrepresentation that occurs when assessments measure skills not taught or not measured skills that are taught (Elliott, Kurtz, Yell, & Tindal, 2017). Given the influence of the CASEL model of SEL on both the revision of the SSIS Classwide Intervention Program and many other intervention programs underway in schools across the globe, having a measure that maximizes construct representation of SEL competencies should contribute to more refined decision making regarding the identification of skills to intervene on and evaluation of intervention outcomes.

Each of the SEL-RFs scales also have an embedded set of 10 items (1 Self-Awareness item, 4 Self-Management items, 1 Social Awareness item, 3 Relationship Skills items, and 1 Responsible Decision-Making item) that make up a Core SEL Scale that is fully aligned with the 10 Core SEL skill units of the SSIS SEL edition Classroom Intervention Program (CIP). **This**

**short, embedded scale that has items for all five CASEL competencies is designed to be used a brief norm-referenced screening and/or progress monitoring scale for explicit use with the SSIS SEL CIP. Table 2 provides a summary of the SSIS SEL Rating Forms' items and internal consistency statistics.**

Teachers and parents indicate the frequency with which students exhibit each social-emotional skill on a 4-point scale of *Never, Seldom, Often, and Almost Always*. Students indicate how true a statement was about each social-emotional skill for them using a 4-point scale of *Not True, A Little True, A Lot True, and Very True*.

The SEL Edition SSIS-Rating Forms are revisions of the Social Skills Improvement System Rating Scales (SSIS-RS). Specifically, the social skill items from the original SSIS were reorganized and assigned to new subscales based on a confirmatory factor analysis that fit the five CASEL domains (Gresham, Elliott, Metallo, Byrd, Erickson, & Altman, in press). The SEL Edition SSIS-Rating Forms were normed on a nationwide representative sample totaling 4,700 children and adolescents ages 3 through 18 years who were assessed in 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 (3-5 years, 5-12 years, 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.

### **Data Analytic Strategies**

The present study examined interrater agreement among the three rater dyads (teacher-parent, teacher-student, and parent-student) with two indices. The first method employed to assess interrater agreement was the calculation of  $q$  correlations among pairs of raters (teacher-

parent, teacher-student, and parent-student). The  $q$  correlations are Pearson correlations between sets of common items by different raters or informants. Specifically, bivariate  $q$  correlations were calculated across raters to examine the convergent and divergent relationships for evidence of interrater reliability and are expressed as Pearson  $r$ . Convergent correlations were the interrater estimates and were those that represented the correlations between informants ratings of the same subscale (e.g., Teacher Self-Awareness—Parent Self-Awareness, Student Self-Management—Teacher Self-**Management, etc.**). **Divergent correlations were those that represented correlations between raters' on different subscales** (e.g., Student Responsible Decision Making—Teacher Social Awareness, Parent Relationship Skills—Student Social Awareness, etc.). Consistent with a multitrait-multimethod matrix (Campbell & Fiske, 1959), one would anticipate the convergent correlations to be higher than the divergent correlations.

When more than one statistical test is performed in analyzing data, some statisticians recommend a more stringent criterion, such as a Bonferroni correction be used for statistical significance than the conventional  $p < .05$ . However, Bonferroni adjustments are often unnecessary and deleterious to sound statistical inference. The biggest problem with this adjustment is that the likelihood of Type II errors is greatly increased thereby creating a scenario that truly important differences are considered non-significant. Describing what tests of significance have been performed and why is usually the best way of dealing with multiple comparisons (Rosenthal & Rubin, 1984; Rothman, 1990). As such, we considered a correlation to be statistically significant if it reached the  $p < .01$  level in the current study.

The second index was an effect size, which captured the extent the dyad agreed about the overall level of social-emotional skills in standard deviation units. Effect sizes were interpreted using Cohen's (1988) conventions for small ( $d = 0.20$ ), medium ( $d = 0.50$ ) and large ( $d = 0.80$ ).

This effect size was calculated by dividing the absolute value of the difference score between raters' standard scores by the pooled standard deviation:  $ES = D / (SD_1 + SD_2) / 2$ .

The effect size captures the extent to which the dyad agreed about the overall level of social emotional skills in standard deviation units. Therefore, effect sizes approaching less than 0.20 indicated high agreement among raters whereas effect sizes reading 0.80 indicated high disagreement among raters. Effect sizes are amenable to arbitrarily assigning the sign or directionality of the effect depending on an interpretation that is consistent with the finding. We assigned positive values to effect sizes to indicate that the adult (teacher or parent) provided less favorable ratings of social-emotional skills. For the teacher-parent dyad, positive values were used to indicate that teachers provided less favorable ratings of social-emotional skills. For example, an effect size of 0.15 representing the agreement between teachers and students on the SSIS-Rating Form Total Social-Emotional Skills would indicate that teachers provided slightly less favorable ratings than did students, however, there was high agreement between their ratings.

For purposes of comparing correlations across raters, scales, and type (convergent and divergent) of correlations, averages were calculated. Because distances between correlations are not equally detectable, the correlations were transformed and placed on a common metric. The Fisher  $Z_r$  was used to transform the correlations and compute an average (see Rosenthal & Rosnow, 2008).

## **Results**

### **Parent-Teacher Agreement**

Both convergent and divergent correlations for parent and teacher ratings on the five SEL domains, SEL Composite, and Core Skills scales can be found in Table 4. The diagonal, bolded

convergent correlations in the table indicate convergent validity. Generally, a significant correlation of  $p < .01$  is indicated by a Pearson's  $r$  coefficient greater than 0.20.

**SEL Composite Scale and Core Skills.** Parent and teachers' ratings on SEL Composite Scale indicated a significant relationship ( $r = 0.34$ ), with ratings on Core Skills also yielding similar results ( $r = 0.38$ ). Comparison of the convergent correlation of the SEL Composite Scale ratings to the mean divergent correlation ( $r = 0.33$ ) revealed no significant difference, indicating a lack of support for the divergent validity of this composite scale. Similar analyses comparing the convergent correlation of the Core Skills to the mean divergent correlation ( $r = 0.35$ ) revealed similar findings, with only a small difference between the correlations.

The standard mean difference effect sizes for the composite scale and Core Skills scale are included in Tables 5 through 8, with gender and scales calculated separately. For scores between informants on the SEL Composite scale, mean difference effect sizes for males ranged from .05 to .13 (Ages 5-12 form) while females ranged .06 to .14 (Ages 5-12 form); however, these effect sizes were small indicating modest agreement between parents and teachers on these ratings. For males, teachers provided slightly less favorable ratings than parents; however, parents provided slightly less favorable ratings for females than teachers.

The mean difference effect sizes for the Core Skills scale for males ranged from .23 to .26 (Ages 8-12 form) while sizes for females ranged from .03 to .07. For females, neither teachers nor parents provided less favorable ratings of core skills than the other. For males, parents and teachers provided less favorable ratings dependent on form, although the small effect size indicates modest agreement between raters.

**SEL Scales.** Convergent correlations on the five SEL domains using the teacher and parent rating forms were all significant, ranging from 0.23 to 0.36 with a median of .33 ( $p < .0001$ ). The highest agreement between parent and teacher ratings was on Self-Management and Responsible Decision Making with the lowest agreement, albeit significant, on the Self-Awareness scale. Comparison of the median convergent correlation ( $r = 0.34$ ) to the mean

divergent correlation ( $r = 0.27$ ) revealed a stronger magnitude than the divergent correlations, although the mean divergent correlation was also significant. Convergent validity coefficients for the teacher form showed that 76% (37/49) of the coefficients were statistically significant ( $p < .001$ ). For parent ratings, 78% (38/49) were statistically significant ( $p < .001$ ). This difference provides moderate support for the convergent validity of the SSIS SEL in that ratings on the same domains by different raters were found to have moderately stronger associations than ratings of different domains by different raters.

### **Parent-Student Agreement**

Both convergent and divergent correlations for parent and student ratings on the five SEL domains, SEL Composite, and Core Skills scales can be found in Table 9. The diagonal, bolded convergent correlations in the table indicate convergent validity.

*SEL Composite Scale and Core Skills.* Parent and student ratings on the SEL Composite Scale demonstrated a significant relationship ( $r = 0.29$ ), indicating high agreement between ratings on this scale. Similar results were found when comparing parent and student ratings on the Core Skills ( $r = 0.26$ ). Comparison of the convergent correlation for the SEL Composite Scale to the mean divergent correlation ( $r = 0.25$ ) revealed a modest magnitude than the divergent correlations, although the mean divergent correlation was also significant. Comparison of the convergent correlation for the Core Skills to the mean divergent correlation ( $r = 0.22$ ) also revealed a slightly stronger magnitude than the divergent correlations, although the mean divergent correlation was also significant.

The standard mean difference effect sizes for the composite scale and Core Skills scale were analyzed separately across gender and forms. For the SEL Composite Scale, the effect size for males across forms ranged from 0.02 to 0.07, with neither parent nor student responding with slightly less favorable ratings. Similar results were found for females across forms, with standard

mean difference effect sizes ranging from 0.03 to 0.07. For the Core Skills, the effect sizes for males across forms ranged from 0.23 to 0.26, indicating a small effect size. For females across forms, mean difference effect sizes ranged from 0.23 to 0.37 (with largest effect sizes on Ages 8-12 Form), revealing small effect sizes in difference in ratings. These results indicate that parents' and students' ratings provided modest agreement in ratings for male and females.

***SEL Scales.*** Convergent correlations on the five SEL domains ranged from a minimum of .17 to a maximum of .27. The highest agreement between parent and student ratings was on Relationship Skills while the lowest agreement was on Self-Awareness, with a correlation approaching significance. Comparison of the median convergent correlation ( $r = 0.24$ ) to the mean divergent correlation ( $r = 0.22$ ) revealed a slightly stronger magnitude than the divergent correlations, although the mean divergent correlation was also significant. This difference provides some support for the convergent validity of the SSIS SEL in that ratings on the same domains by different raters were found to have more modest associations than ratings of different domains by different raters. For the student rating form, 69% (34/49) of the convergent validity coefficients were statistically significant ( $p < .0001$ ).

### **Teacher-Student Agreement**

Both convergent and divergent correlations for teacher and student ratings on the five SEL domains, SEL Composite, and Core Skills scales can be found in Table 10. The diagonal, bolded convergent correlations in the table indicate convergent validity.

***SEL Composite Scale and Core Skills.*** Teacher and student ratings on the SEL Composite Scale demonstrated a significant relationship ( $r = 0.27$ ), indicating high agreement between ratings on this scale. Similar results were found when comparing teacher and student ratings on Core Skills ( $r = 0.30$ ). Comparison of the convergent correlation for the SEL



Composite Scale to the mean divergent correlation ( $r = 0.23$ ) revealed a more modest magnitude than the divergent correlations, although the mean divergent correlation was also significant. Comparison of the convergent correlation for the Core Skills to the mean divergent correlation ( $r = 0.23$ ) also revealed a stronger magnitude than the divergent correlations, although the mean divergent correlation was also significant.

Standard mean difference effect sizes for the composite scale and Core Skills scale were analyzed separately across gender and forms. For males the mean difference effect sizes between teacher and student ratings on the SEL Composite Scale were small, ranging from .02 to .06. Similar results were found for females, with effect sizes ranging from .01 to .07, with teachers providing slightly less favorable responses than males and with females providing slightly less favorable responses than teachers. However, these effect sizes are small, indicating strong agreement between informants. For males the mean difference effect sizes between teacher and student ratings on Core Skills ranged from a small effect size of .14 and from a medium effect size of .41 (Ages 8-12 form), with teachers responding with less favorable ratings. For females, effect sizes ranged from .23 to .39 (Ages 13-18 form), with females responding with slightly less favorable responses on the Ages 13-18 form. Effect sizes approaching .50 indicate some moderate differences between informants.

***SEL Scales.*** Convergent correlations between teacher and student ratings on the five SEL domains ranged from 0.14 to 0.26. The highest agreement between teacher and student ratings was on both Relationship Skills and Responsible Decision Making. The lowest agreement was on Social Awareness, with a nonsignificant correlation. Comparison of the median convergent correlation ( $r = 0.23$ ) to the mean divergent correlation ( $r = 0.20$ ) revealed a slightly stronger magnitude than the divergent correlations, although the mean divergent correlation was also

significant. **This difference provides moderate support for the convergent validity of the SSIS SEL in that ratings on the same domains by different raters were found to have modest associations than ratings of different domains by different raters.**

### **Comparison of Similar and Dissimilar Informants**

The results of the analysis of interrater reliability between similar informants (teacher-teacher and parent-parent) and dissimilar informants (teacher-parent) can be found in Table 11. Teacher-teacher and parent-parent dyads demonstrated stronger agreement than did teacher-parent dyads. Across all SEL domains and composite scores, similar informants showed higher interrater reliability estimates than dissimilar. The average interrater correlations for teacher-teacher, parent-parent, and teacher-parent were 0.63, 0.57, and 0.33, respectively.

As indicated in Table 12, the highest convergent validity estimates by domain and rater revealed strongest agreement across various SEL scales and composite scales while the weakest agreement was more consistently observed on the Self-Awareness or Social Awareness SEL domain scales.

### **Discussion**

Researchers and practitioners alike have been encouraged for decades to use multiple sources of information when assessing children's social behavior to capture the likelihood of its variability across situations and settings (e.g., Achenbach, McConaughy, & Howell, 1987; De Los Reyes & Kazdin, 2005; Ruffalo & Elliott, 1997). The current study was conducted with this assessment tenet in mind and with the goal of replicating an investigation of cross-informant agreement with SSIS ratings of children's social skills (i.e., Gresham et al., 2010). In this study, we also extended the focus of raters' agreement to children's social emotional skills; in particular the SEL competency domains of self-awareness, self-management, social awareness, relationship

skills, and responsible decision making as assessed by the SSIS SEL Edition Rating Forms. Very few measures of these constructs actually exist, yet they are the focus of many school-based intervention programs operating in the United States and beyond (CASEL, 2015).

### **Key Findings**

Three predictions were posited based on our review of cross-informant research in general and specifically previous findings with the social skills version of the SSIS. Using representative subsamples of male and female students from grades 3 to 12, and their parents and teachers, we tested the predictions and found modest to strong support for each of them. Specifically, we found that (a) pairs of different types of informants (teacher-parent, teacher-student, and parent-student) exhibited greater than chance levels of agreements as indexed by significant interrater correlations, (b) teacher-parent informants showed higher correlations (on average .33) than teacher-student (average .23) and parent-student (average .24) pairs across all SEL competency domains assessed, and (c) pairs of similar informants (teacher-teacher average .63 and parent-parent average .57) exhibited significantly higher correlations than pairs of dissimilar informants (average of .27 across teacher-parent, teacher-student, parent-student pairs).

The comparison of SEL Composite Scale scores between rating pairs for males and females across the age spectrum provided an alternative means of examining the agreement between dissimilar raters of children's social competence at a global level. Specifically, the results of composite score differences between raters yielded consistently low effect sizes; thus indicating low disagreement or rather a high overall level of agreement in ratings for the SSIS SEL Rating Forms. Interestingly, when the ratings by the same rater pairs were compared for the Core Skills Scale, a subset of 10 basic SEL skills that cut across each of the five SEL

competency domains, the effect sizes were substantially higher, indicating that the raters' agreement was only moderate. These findings with different length composite score scales and the use of an effect size index add new information and methods to the cross-informant agreement research literature.

### **Comparison of Findings to Previous Cross-Informant Research**

Our findings were very consistent with those reported by Gresham et al. (2010), although most agreement correlations were slightly higher for the SEL competency domains, than the social skills scales of the SSIS. These results are likely due to the fact that each SSIS SEL Competency Scale has a few more items than each Social Skills Scale, thus better representing the constructs they are designed to measure. Our results with the SSIS SEL cross-informant agreement analyses also were highly consistent with Renk and Phares' (2004) meta-analysis of cross-informant ratings of children and adolescents' social competence. Specifically, these researchers reported effect size estimates for 74 studies with a mean of .38 for parent-teacher pairs, a mean of .25 for teacher-student pairs, and a mean of .21 for parent-student pairs. Thus, the current investigation indicated that teachers, parents, and students do not seem to perceive levels of social emotional skills very differently, and this conclusion is consistent with a number of previous studies that have investigated an array of rating scale assessments of children's social behaviors across a variety of settings.

In addition to our specific findings about the magnitude of agreements across informants for the same SSIS social emotional scales, we learned about informants' ratings of different SEL social emotional scales. That is, applying a multitrait-multisource logic to examine the correlations among ratings derived from different raters can be used to assess the convergent validity coefficients of the SSIS SEL scales (Campbell & Fiske, 1959). Across the different rater

pairings, our results indicated that the convergent validity coefficients (on the diagonal) were consistently higher than the divergent/discriminant validity coefficients (off the diagonal). The differences, however between the convergent and divergent validity coefficients were relatively small across raters and scales.

Overall, the evidence indicates that discrepancies among informants' reports reflect setting-based differences among informants' opportunities for observing behavior (Achenbach et al., 1987; De Los Reyes & Kazdin, 2005). However, it is unclear as to whether informants consistently use setting information when making judgments about behavior. This is an important issue because one might assume that parent reports represent "home" behaviors and teacher reports represent "school" behaviors. There is some evidence regarding whether informants consistently use setting information when making reports of behavior. For example, an experimental study using a sample of experienced clinicians who read vignettes describing the home, school, and peer settings of children expressing symptoms of conduct disorder showed that these clinicians varied widely on when they applied setting information to their clinical judgments (De Los Reyes & Marsh, 2011). It appears that clinicians use setting information to make clinical judgments, but do so inconsistently. This is an important avenue for future research on the use of multiple informants' judgments of social-emotional skills.

### **Limitations and Future Research**

This study had several design features important for fair, meaningful cross-informant comparisons. Specifically, the samples of raters were from a large, representative national standardization sample; the items on the various teacher, parent, and student forms were the same; there was no missing data; and raters were naïve to the purpose of the study. Like any study, however, there were notable limitations. First, the item rating anchors for the student

form concerned how true a statement was about a particular SEL skill, whereas the item rating anchors used by teachers and parents concerned the frequency of the SEL skill. Second, the ratings were not completed on the same day or in the same location by the various raters; teachers and students completed the ratings during the same week at their school, however, parents completed their ratings within the same month and most likely in their homes. Future cross-informant investigators are encouraged to have all ratings done at the same time, in the same place, and with the same item rating anchors to control these potential sources of variance. Admittedly, the variance accounted for by these design variations is likely small, but given the interest in using multiple ratings and sources of evidence when assessing children's social behavior, it is important to continue to refine research designs and the assessments used.

### **Implications for Practice**

Users of multi-rater assessments like the SSIS Rating Scales or SSIS SEL Edition Rating Forms can expect to find some disagreement across raters because humans' social behaviors is influenced by social situations and environments. That is, teachers and parents, teachers and students, or parents and students can be expected to rate the frequency or presences of some social emotional skills somewhat differently; however, on balance, the present research indicated much more agreement than disagreement, and when focusing on composite scores, the agreement levels are relatively high. Rating scales like the SSIS SEL Edition can be supplemented by direct observations and interviews to facilitate understanding of possible disagreements among raters and to provide more qualitative evidence about children's social emotional skills.

**Overall, there are several unanswered practical questions with regard to informant discrepancies of social emotional learning skills that impact everyday educational practice.**

**For example, how are educators supposed to interpret informant discrepancies when they encounter them in their work? Does the presence of informant discrepancies impact educators' decisions regarding student response to intervention, classification, and social emotional learning interventions? Many have dismissed informant discrepancies as representing random error due to informants' biased perspectives, unreliable reporting, or lack of awareness, whereas recent data indicate that informant discrepancies are informative and clinically and educationally useful (Achenbach, 2006; De Los Reyes & Kazdin, 2005). Future research should establish guidelines to interpret informant discrepancies and use this information for educational decision-making regarding social emotional learning skills.**

### **Conclusions**

This study was one of the first to examine cross-informant agreement regarding children's SEL skills consistent with the CASEL model of competencies. The magnitude of the agreement evidence resulting from the current study was substantial and quite positive in the context of previous cross-informant rating studies. Empirical evidence in support of three fundamental predictions regarding rater agreement was advanced, and the relevance of an effect size index for cross-informant research demonstrated. Collectively, this evidence indicated the SSIS SEL Edition Rating Forms functioned very similar to their predecessors within the SSIS Rating Scales with regard to the multi-rater assessment of children's social behavior, and in particular, when assessing key SEL competencies that are part of numerous programs designed to improve children's social emotional wellbeing across school and home environments. In summary, the cross-informant agreement evidence contributes new information regarding the

psychometric quality of the SSIS SEL Rating Forms and facilitates more multi-rater assessments of students in SEL programs.

### References

Achenbach, T. (2006). As others see us: Clinical research implications for psychopathology.

*Current Directions in Psychological Science, 15*, 94-98.

Achenbach, T., & Rescorla, L. (2001). *Manual for the ASEBA school-age forms and profiles*.

Burlington, VT: University of Vermont Research Center Children, Youth, and Families.

Achenbach, T., McConaughy, S., & Howell, C. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-232.

Campbell, D., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Caprara, G., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P. (2000). Prosocial foundations of children's academic achievement. *Psychological Science, 11*, 302-306.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). Mahwah, NJ: Erlbaum.

Collaborative for Academic, Social, and Emotional Learning (2015). *District guide to systemic social and emotional learning*. Chicago: Author.

Comer, J., & Kendall, P. (2004). A symptom-level examination of parent-child agreement in the diagnosis of anxious youth. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*, 878-886.

De Los Reyes, A., & Kazdin, A. (2005). Informant discrepancies in assessment of childhood psychopathology: A critical review. *Psychological Bulletin, 131*, 483-509.



- DiPerma, J., & Elliott, S.N. (2002). Promoting academic enablers to improve student achievement. *School Psychology Review, 31*, 293-298.
- Durlak, J. A., Weissberg, R.P., Dymnicki, A.B., Taylor, R.D., & Schellinger, K.G. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*, 405-432.
- Gresham, F.M. (2005). Response to intervention: An alternative means of identifying students as emotionally disturbed. *Education and Treatment of Children, 28*, 328-344.
- Gresham, F.M., & Elliott, S.N. (1990). *Social Skills Rating System*. Minneapolis, MN: Pearson Assessments.
- Gresham, F.M., & Elliott, S.N. (2008). *Social Skills Improvement System Rating Scales*. Minneapolis, MN: Pearson Assessments.
- Gresham, F.M., & Elliott, S.N. (2017). *Social Skills Improvement System Social Emotional Learning Edition Rating Forms*. Minneapolis, MN: Pearson Assessments.
- Gresham, F.M., Elliott, S.N., Cook, C.R., Vance, M.J., & Kettler, R. (2010). Cross-informant agreement for ratings for Social Skill and Problem Behavior Ratings: An investigation of the Social Skills Improvement System –Rating Scales. *Psychological Assessment, 22*, 157-166.
- Gresham, F. M., Elliott, S.N., Metallo, S., Byrd, S., Wilson, E., Erickson, M., & Altman, R. (in press). Comparability of SSIS Rating Scales to the SSIS Social Emotional Learning Rating Forms: Content and psychometric comparisons across elementary and secondary levels. *Journal of School Psychology*.

- Humphrey, N., Kalamouka, S., Wigelsworth, M. Lendrum, A. , Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and Psychological Measurement, 71*, 617-637.
- Kamphaus, R., & Reynolds, C.R. (2016). *Behavioral Assessment System for Children-3*. Minneapolis, MN: Pearson Assessments.
- Malecki, C., & Elliott, S.N. (2020). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17*, 1-23.
- McCleod, R., McNamme, J, Boyle, M., Offord, D., & Friderich, M. (1999). Identification of childhood psychiatric disorder by informant: Comparisons of clinic and community samples. *The Canadian Journal of Psychiatry, 44*, 144-150.
- McKown, C. (2017). Social-emotional assessment, performance, and standards. *The Future of Children, 27(1)*, 157-177.
- MTA (Multimodal Treatment Study of Children with ADHD) Cooperative Group (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry, 56*. 1073-1086.
- Offord, D., Boyle, M., Racine, Y., Szatmari, P., Fleming, J., Sanford, M.,& Lipman, E. (1996). Integration assessment data from multiple informants. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*, 1078-1085.
- Rothman, K.J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology, 1*, 43-46.
- Sparrow, S., Cichetti, D., & Balla, D. (2005). *Vineland Adaptive Behavior Scales* (2<sup>nd</sup> ed.). Minneapolis, MN: Pearson Assessments.

- Taylor, R.D., Oberle, E., Durlak, J.A., & Weissberg, R.P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development, 88*(4), 1156-1171.
- U.S. Census Bureau (2006). *Current population survey, March 2006*. (Machine-readable data file). Washington, DC: Author.
- Valo, S., & Tannock, R. (2010). Diagnostic instability of DSM-IV ADHD subtypes: Effects of informant source, instrumentation, and methods of combining symptom reports. *Journal of Clinical Child and Adolescent Psychology, 39*, 749-760.
- Weissburg, R.P., Durlak, J.A., Domitrovich, C.E., & Gullotta, T.P. (2015). Social and emotional learning: Past, present, and future. In J. Durlak, C. Domitrovich, R. Weissburg, & T. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 3-19). New York: Guilford Press.
- Wentzel, K. (1993). Does being good make the grade? Social behavior and academic competence in middle school. *Journal of Educational Psychology, 85*, 357-364.

Table 1

*Descriptions of SSIS SEL Scale Items and Related Statistics*

Scale	Example Items	$\alpha^*$				
		Teacher	Parent	Parent-Spanish Version	Student	Student-Spanish Version
Self-Awareness	Acts lonely. Asks for help from adults.	0.80	0.77	0.74	0.80	0.82
Self-Management	Has temper tantrums. Stays calm when teased.	0.92	0.88	0.86	0.83	0.86
Social Awareness	Forgives others. Tries to comfort others.	0.90	0.87	0.85	0.80	0.81
Relationship Skills	Takes turns in conversations. Interacts well with other children.	0.91	0.88	0.88	0.87	0.89
Responsible Decision Making	Respects the property of others. Is well-behaved when unsupervised.	0.82	0.84	0.84	0.74	0.77
Academic Competence		0.97**	***	***	***	***
SEL Composite Scale		0.97	0.96	0.95	0.95	0.96
Core Skills		0.91	0.85	0.84	0.82	0.86

\*Mean coefficient alpha across all age levels, combined gender norms

\*\*Mean score includes only Ages 5-12 and Ages 13-18 reliability norm

\*\*\*Academic Competence not included in scale

Table 2

*Demographic Characteristics for Subsample With Dissimilar Raters*

Variable	Frequency	%
Gender		
Male	106	63
Female	62	37
Grade		
3rd	28	17
4th	19	11
5th	27	16
6th	27	16
7th	16	10
8th	12	7
9th	7	4
10th	5	5
11th	14	8
12th	13	8
Total	168	100

Table 3.

*Demographic Characteristics of Interrater Reliability Subsample with Similar Raters*

Characteristic	Teacher Form	Parent Form
<i>N</i>	54	110
Age		
<i>M</i> (years:months)	8:7	9:0
<i>SD</i> (months)	40.1	49.6
Gender		
Female	32	44
Male	22	66
Race/ethnicity		
African American	11	10
Hispanic	12	8
White	25	87
Other	6	5
Mother's education		
Grade 11 or less	7	5
Grade 12 or GED	24	28
1-3 years of college	13	46
4 or more years of college	10	31

*Note.* GED = General Equivalency Diploma

Table 4

*Teacher and Parent Correlations for SEL Scales, Composite Scale and Core Skills*

Parent	Teacher						
	1	2	3	4	5	6	7
1. Self-Awareness	<b>0.23*</b>	0.16	0.15	0.21*	0.16	0.21*	0.18
2. Self-Management	0.24*	<b>0.36*</b>	0.28*	0.34*	0.36*	0.36*	0.36*
3. Social Awareness	0.17	0.25*	<b>0.28*</b>	0.30*	0.26*	0.25*	0.25*
4. Relationship Skills	0.26*	0.25*	0.27*	<b>0.34*</b>	0.29*	0.27*	0.27*
5. Responsible Decision Making	0.23*	0.33*	0.29*	0.33*	<b>0.36*</b>	0.34*	0.34*
6. SEL Composite Scale	0.27*	0.32*	0.31*	0.37*	0.35*	<b>0.34*</b>	0.34*
7. Core Skills	0.27*	0.37*	0.31*	0.36*	0.38*	0.38*	<b>0.38*</b>

*Note.* Boldface indicates convergent validity coefficients.

\* $p < .05$ .

Table 5  
*Standardized Mean Difference Effect Size for Different Between Raters' Score on the SEL Composite Scale - Males*

Rater dyad	<i>M</i> SS <sub>1</sub>	<i>SD</i> <sub>1</sub>	<i>M</i> SS <sub>2</sub>	<i>SD</i> <sub>2</sub>	<i>D</i>	ES
Parent-Teacher						
Ages 3-5	488.2	66.8	484.8	68.9	3.4	0.05
Ages 5-12	494.2	61.7	485.8	67	8.4	0.13
Ages 13-18	495.1	63.6	492.7	60.2	2.4	0.04
Parent-Student						
Ages 8-12	494.2	61.7	489.9	67.6	4.3	0.07
Ages 13-18	495.1	63.6	494.1	64.2	1	0.02
Teacher-Student						
Ages 8-12	485.8	67	489.9	67.6	4.1	0.06
Ages 13-18	492.7	60.2	494.1	64.2	1.4	0.02

*Note.* SS = standard score |*D*| = absolute difference score; ES = effect size.

<sup>1</sup>Mean standard score and standard deviation for first rater. <sup>2</sup>Mean standard score and standard deviation for second rater.



Table 6  
*Standardized Mean Difference Effect Size for Different Between Raters' Score on the SEL Composite Scale - Females*

Rater dyad	<i>M</i> <i>SS</i> <sub>1</sub>	<i>SD</i> <sub>1</sub>	<i>M</i> <i>SS</i> <sub>2</sub>	<i>SD</i> <sub>2</sub>	<i>D</i>	ES
Parent-Teacher						
Ages 3-5	512.3	53.7	515.4	56.5	3.1	0.06
Ages 5-12	505.8	59.4	514	57.6	8.2	0.14
Ages 13-18	504.1	59.8	506.8	67.5	2.7	0.04
Parent-Student						
Ages 8-12	505.8	59.4	509.8	59.4	4	0.07
Ages 13-18	504.1	59.8	506	65.3	1.9	0.03
Teacher-Student						
Ages 8-12	514	57.6	509.8	59.4	4.2	0.07
Ages 13-18	506.8	67.5	506	65.3	0.8	0.01

*Note.* *SS* = standard score; |*D*| = absolute difference score; ES = effect size.

<sup>1</sup>Mean standard score and standard deviation for first rater. <sup>2</sup>Mean standard score and standard deviation for second rater.

Table 7  
*Standardized Mean Difference Effect Size for Different Between Raters' Score on Core Skills - Males*

Rater dyad	<i>M</i> SS <sub>1</sub>	<i>SD</i> <sub>1</sub>	<i>M</i> SS <sub>2</sub>	<i>SD</i> <sub>2</sub>	<i>D</i>	ES
Parent-Teacher						
Ages 3-5	18.8	5	20.5	5.4	1.7	0.33
Ages 5-12	20.8	4.4	19.9	5.6	0.9	0.18
Ages 13-18	21.5	4.7	21.1	4.9	0.4	0.08
Parent-Student						
Ages 8-12	20.8	4.4	22	4.7	1.2	0.26
Ages 13-18	21.5	4.7	20.4	4.8	1.1	0.23
Teacher-Student						
Ages 8-12	19.9	5.6	22	4.7	2.1	0.41
Ages 13-18	21.1	4.9	20.4	4.8	0.7	0.14

*Note.* SS = standard score; |D| = absolute difference score; ES = effect size.

<sup>1</sup>Mean standard score and standard deviation for first rater. <sup>2</sup>Mean standard score and standard deviation for second rater.

Table 8  
*Standardized Mean Difference Effect Size for Different Between Raters' Score on Core Skills - Females*

Rater dyad	<i>M</i> SS <sub>1</sub>	<i>SD</i> <sub>1</sub>	<i>M</i> SS <sub>2</sub>	<i>SD</i> <sub>2</sub>	<i>D</i>	ES
Parent-Teacher						
Ages 3-5	20.4	4.3	22.7	4.7	2.3	0.51
Ages 5-12	21.6	4.2	22.1	5	0.5	0.2
Ages 13-18	22.3	4.4	23.2	5.1	0.9	0.19
Parent-Student						
Ages 8-12	21.6	4.2	23.2	4.4	1.6	0.37
Ages 13-18	22.3	4.4	21.2	5.2	1.1	0.23
Teacher-Student						
Ages 8-12	22.1	5	23.2	4.4	1.1	0.23
Ages 13-18	23.2	5.1	21.2	5.2	2	0.39

*Note.* SS = standard score; |*D*| = absolute difference score; ES = effect size.

<sup>1</sup>Mean standard score and standard deviation for first rater. <sup>2</sup>Mean standard score and standard deviation for second rater.

Table 9  
*Parent and Student Correlations for SEL Scales, Composite Scale and Core Skills*

Student	Parent						
	1	2	3	4	5	6	7
1. Self-Awareness	<b>0.17</b>	0.23*	0.24*	0.24*	0.22*	0.27*	0.25*
2. Self-Management	0.18	<b>0.24*</b>	0.21*	0.21*	0.19	0.25*	0.25*
3. Social Awareness	0.11	0.17	<b>0.25*</b>	0.24*	0.15	0.23*	0.20*
4. Relationship Skills	0.22*	0.22*	0.24*	<b>0.27*</b>	0.19	0.28*	0.23*
5. Responsible Decision Making	0.16	0.23*	0.19	0.20*	<b>0.23*</b>	0.25*	0.26*
6. SEL Composite Scale	0.19	0.25*	0.26*	0.27*	0.23*	<b>0.29*</b>	0.27*
7. Core Skills	0.18	0.24*	0.21*	0.22*	0.21*	0.26*	<b>0.26*</b>

*Note.* Boldface indicates convergent validity coefficients.

\* $p < .05$ .

Table 10  
*Teacher and Student Correlations for SEL Scales, Composite Scale and Core Skills*

Student	Teacher						
	1	2	3	4	5	6	7
1. Self-Awareness	<b>0.16</b>	0.22*	0.14	0.25*	0.22*	0.23*	0.23*
2. Self-Management	0.15	<b>0.23*</b>	0.09	0.16	0.21*	0.20*	0.23*
3. Social Awareness	0.19	0.20*	<b>0.14</b>	0.20*	0.14	0.20*	0.18
4. Relationship Skills	0.24*	0.22*	0.13	<b>0.26*</b>	0.19	0.25*	0.22*
5. Responsible Decision Making	0.17	0.27*	0.16	0.24*	<b>0.26*</b>	0.26*	0.29*
6. SEL Composite Scale	0.21*	0.27*	0.15	0.26*	0.24*	<b>0.27*</b>	0.26*
7. Core Skills	0.18	0.28*	0.13	0.24*	0.26*	0.26*	<b>0.30*</b>

*Note.* Boldface indicates convergent validity coefficients.

\* $p < .05$ .

Table 11  
*Interrater Reliability Correlations Between Similar and Dissimilar Informants*

Scale	Rater dyad		
	Teacher-Teacher	Parent-Parent	Teacher-Parent
1. Self-Awareness	0.53	0.34	0.23
2. Self-Management	0.67	0.63	0.36
3. Social Awareness	0.56	0.53	0.28
4. Relationship Skills	0.72	0.6	0.34
5. Responsible Decision Making	0.51	0.59	0.36
6. Academic Competence	0.62	***	***
7. SEL Composite Scale	0.69	0.62	0.34
8. Core Skills	0.7	0.68	0.38
<b>Average rater dyad r</b>	<b>0.63</b>	<b>0.57</b>	<b>0.33</b>

\*\*\* Academic Competence not included in scale

Table 12

*Comparison of SEL Scales With Strongest and Weakest Agreement*

Raters	SEL interrater reliability	
	Strongest	Weakest
Teacher-Parent	Self-Management, Responsible Decision Making, Core Skills	Self-Awareness
Teacher-Student	Relationship Skills, Responsible Decision Making, Core Skills	Social Awareness
Parent-Student	Relationship Skills, SEL Composite Scale	Self-Awareness