The Assessment of Children’s Social Skills Through Self-Report: A Potential Screening Instrument for Classroom Use

Carla Kmett Danielson and Carolyn Roecker Phelps

The Children’s Self-Report Social Skills Scale (CS), a 21-item instrument, was developed to measure children’s perspectives on their own social skills. Test–retest reliability and internal consistency reliability of CS scores were .74 and .96, respectively. Principal component analysis revealed 3 reliable components: Social Rules, Likeability, and Social Ingenuousness.

Traditionally, parents and teachers have been among the most common sources of information regarding a child’s behavior. Although these adult ratings can offer useful information, the accuracy of their reports can possibly be distorted by factors such as reporting biases (e.g., “middle-class” bias) and depression (Youngstrom, Izard, & Ackerman, 1999; Youngstrom, Loeber, & Stouthamer-Loeber, 2000). The potential biases in these reports can be viewed as quite costly because important decisions are often based on these assessments. Thus, assessment decisions should not necessarily rely solely on the adult’s report of the child’s behavior. Rather, researchers and clinicians have emphasized the importance of gathering a self-report from the child so that an additional source of information is available and the child is able to convey his or her own perceptions and behaviors (Achenbach, 1995; La Greca, 1990; Loeber, Green, & Lahey, 1990).

A clear advantage of a child’s self-report is that he or she is in the unique position to report on behaviors across different situations, including home, classroom, playground, sports practice, and so forth. Beitchman and Corradini (1988) stated,

The limitations of teacher and parent rating scales point to the necessity of developing similar instruments that can better represent the child’s point of view. Because the child is, of course, common to both school and home, it may be that his view of himself is the most important predictor of his behavior. (p. 478)

Advantages of child self-report measures include the following: (a) The instruments are generally inexpensive and can be easily administered in a wide variety of settings, such as an outpatient clinic, a pediatrician’s office, a school, or home (Beitchman & Corradini, 1988), and (b) meaningful information—the child’s perceptions and cognitions—is provided that is not otherwise accessible to other reporters.

Because of these advantages, several children’s self-report scales are being used both clinically and in research. These measures have been developed to assist in the assessment of several domains, such as depression (e.g., Birleson, 1981; Kovacs, 1981), coping skills (Causey & Dubow, 1992), and overall functioning (e.g., Achenbach, 1991a). One domain that is lacking in the area of self-report measures is social skills. Social skills are neces-
sary for all children, so that the child is able to interact within his or her social milieu. Furthermore, social skill deficits can prove detrimental to a child's later functioning, highlighting the importance of assessment in this domain (Dodge & Richard, 1985; Elliott & Gresham, 1987). In this article, we present a measure that was developed for assessment in the area of children's social skills. First, we provide a brief definition of social skills and explain the importance and need for such a measure.

Many conceptualizations of social skills proposed by researchers fall somewhere on a continuum between two basic models, the trait model of social skills and the molecular model of social skills (McFall, 1982). The trait model treats social skills as an underlying personality characteristic, whereas the molecular model defines social skills in terms of specific, observable units of behavior, laying the foundation for an individual's overall performance in each social situation. Dissatisfied with the two prevailing models, McFall proposed defining social skills as specific abilities that enable a person to perform competently at particular tasks. We purport that it is more important to focus on the utilization of a social task rather than the mere existence of the ability to perform that social task. In other words, does a child have a social skill if he or she has the ability to initiate play with other children, but almost never uses this ability? Furthermore, assessment efforts should be targeted toward measuring whether a child is actually performing these necessary skills. Thus, we define social skills as interactions and tasks that a child performs competently within his or her social milieu.

A growing body of literature indicates that the development of adequate social skills during childhood has important and far-reaching ramifications (Merrell, 1993). Social skills are related to many areas of functioning in a child's life, such as good treatment and respect from parents (Putallaz & Heflin, 1990), high self-esteem (Boivin & Begin, 1989), and overall happy quality of life (Rubin, Booth, Rose-Krasnor, & Mills, 1995). Similarly, deficits in the development of adequate social skills can have detrimental effects on a child's current and future level of functioning. A variety of clinical disorders and problems have also been associated with social skill deficits and peer rejection, both in childhood and early adulthood (e.g., Cowen, Pederson, Barbigian, Izzo, & Trost, 1973; Loeb, 1985; Parker & Asher, 1987). In addition, researchers have found that children who manifest social skills deficits are likely to evidence lower levels of academic achievement and to score below average on measures of cognitive and emotional development (Green, Forehand, Beck, & Vosk, 1980; Hubbard & Cole, 1994).

The empirical demonstration that the development of children's social skills is critical to their overall functioning highlights the importance of the identification and measurement of social skills. A variety of assessment methods have been used to evaluate the social skills of children. These methods include direct observation, behavioral interviews, and peer evaluations. Psychometric checklists and rating forms completed by caregivers and teachers are another popular method of assessment of children's social skills (Elliott & Gresham, 1987). Although all of these aforementioned methods have been the most common forms of assessment for children's social skills, these techniques can be time consuming, and they rely on the inference of the observer (e.g., the teacher must infer whether a child believes that he or she has no friends). Despite emphasis on the importance of gathering a self-report from the child to counter such disadvantages, few children's self-report measures that are specific to the construct of social skills exist, especially in comparison with the number of other constructs that are frequently measured in children by self-report (e.g., depression). Two of the existing measures of social skills are the Matson Evaluation of Social Skills for Youngsters (MESSY; Matson, Rotatori, & Helsel, 1983) and the List of Social Situation Problems (LSSP; Spence, 1980). The MESSY is a 62-item instrument in which children rate themselves on both positive and negative social behaviors. Although psychometric properties have been reported for scores on the measure, later factor analyses failed to...
replicate the same underlying factors as reported by the authors of the MESSY (Spence & Little, 1990). In addition, screening instruments shorter than 62 items may be more useful in situations where all students within a classroom, grade, or school are tested simultaneously, so that the screening can be administered quickly without losing the attention of the children. In addition, some of the MESSY items are very similar to one another (e.g., Item 15: "I feel angry or jealous when someone else does well"; Item 54: "I am jealous of other people") and may be unnecessarily redundant.

The 60 items of the LSSP focus on identifying the specific interpersonal situations that are experienced as difficult, uncomfortable, or hard to handle in order to help guide treatment in the context of social skills training programs. Again, psychometric properties of the scores from the measure have been reported. However, one reason the LSSP may not be optimal as a brief screening self-report measure of children's social skills is that some of the items of the LSSP may be more specific to other constructs than social skills. Specifically, the largest subscale of the LSSP reported by Spence and Little (1990) was Social Anxiety/Assertiveness, accounting for 31% of the variance in the factor analysis. Many of the 12 items on this subscale seemed to be assessing affective symptoms of social anxiety problems rather than general behavioral social skill deficits (e.g., Item 52: "[Do you] often worry that you might make a fool of yourself in front of other people?"). Another reason that the LSSP may not be optimal is that the language in some of the items may be confusing to children in the United States because it was developed in Australia and contains phrases uncommon to U.S. culture (e.g., Item 48: "[Do you] have trouble getting on with your parents?").

Two uses of child self-report are salient in current school and clinical assessment practices. First, self-report plays a role in the recent movement in the field toward gathering data on the child's behavior from multiple informants (Kolko & Kazdin, 1993). This multiple-informant approach to assessment considers the potential benefit of including all sources of data to form the most complete picture of the child. In using this approach, much research is currently focusing on optimizing methods of combining information from different reporters, through techniques such as algorithmic approaches and q correlations (e.g., Achenbach, 1991a; Piacentini, Cohen, & Cohen, 1992; Youngstrom et al., 2000). Therefore, it is important to have measures that allow each reporter, including the child, to contribute his or her piece in a meaningful and practical fashion.

Second, self-report measures can be used as screening instruments in a multiple-gating approach to assessment for the purposes of early identification and primary intervention. That is, because all children are involved in social interactions and because social skills deficits potentially can have a critical impact on a child's future (e.g., Cowen et al., 1973), it may be beneficial to identify those children who are experiencing social difficulties sooner rather than later. Although gathering information from multiple informants and conducting extensive clinical interviews with the child are the most exhaustive way to identify children with social skill problems, time and financial barriers make it impractical to screen all children via this process. In searching for a way to screen a large number of children in the most time- and cost-effective way, Reynolds (1994) has described a procedure, referred to as multiple gating, which may help with the identification of problems in children. This procedure is modeled from a multistage assessment for the screening and identification of depression in children and adolescents in a school setting. The procedure includes a three-stage process, which begins with the administration of a brief screening device to a large group of students. Those children who exceed a determined cutoff score are retested. When meeting the specified screening score a second time (i.e., demonstrating clinical levels of depression at Stages 1 and 2), the children then participate in an individual clinical interview. This process has been validated in its use in screening for depression in children (Kahn, Kehle, & Jenson, 1988; Reynolds & Evert, 1991) and can be used to inform the target of treatment programs. For example, Seligman (1995) de-
scribed a school-based prevention program in which children who were screened at school for depressive symptoms were taught to be more optimistic. This process can be generalized to the screening of other problems, such as social skill deficits, in children. A brief screening device that can quickly and easily be administered in a school setting is necessary for the first stage of this process. In considering these uses of self-report and the limitations of the aforementioned self-report measures, the purpose of this study was to develop a brief self-report measure, the Children's Self-Report Social Skills Scale (CS4), for the assessment of children's social skills.

As part of the development of the CS4, which is explained in detail in the Methods section, the validity of the scores on the proposed scale is examined by first examining how well the scores from the instrument agree with the scores of another measure of the children's social skills: peer nomination scores. Peer nomination scores have been recognized as valid, and at times preferable, means of assessment of peer acceptance (Parker & Asher, 1987) as well as children's social skills (Cowen et al., 1973; La Greca, 1990). The present study uses a modification of the Peer Evaluation Inventory (PEI; Pekarik, Prinz, Liebert, Weintraub, & Neale, 1976), which is also explained in more detail in the following section.

To further examine the validity of the scores of the CS4, we wanted to examine the relationship between the children's social skills scores and the scores on a measure of depression. This procedure of assessing the association between the social skills scores and depression scores is consistent with some of the validation procedures that occurred with the aforementioned social skill measures in existence, the MESSY and the LSSP (Spence & Little, 1990). A relatively strong, negative association between reported depressive symptoms and social skills behavior has been demonstrated in the literature (Friedt, 1990; Kazdin, 1990). That is, children who are depressed generally exhibit social skills deficits (Kovacs, 1989). For example, in experimental situations, depressed children have performed less skillfully on interpersonal tasks and are preferred less often as playmates in social situations (e.g., Altmann & Gotlib, 1988; Sacco & Graves, 1984). This relationship between poor social skills and depression is particularly important to consider when determining treatment options for the child. According to the present study, a child may have the ability to engage in certain social behaviors but may not be using those skills for various reasons, for example, if the child is shy, lacks self-confidence, or is depressed. When considering the techniques that are involved in improving social skills and mood (i.e., decreasing depression), considerable overlap may be found. For example, teaching the child to be assertive, to learn conflict resolution, and to increase the frequency of positive, pleasant social events with peers are strategies applied in social skills and depression interventions. Thus, although the purposes of including the depression measure in the present study focused on establishing construct validity, it is also important to consider its treatment implications in the context of social skills interventions.

In sum, social skills have been demonstrated to play a pivotal role in a child's current and future functioning. An accurate and appropriate means to assess children for social skill deficits is imperative, so that the children's own perspective is provided and children who are experiencing social skill difficulties are identified. Self-report social skill scales that have been developed are not optimal for use in a screening process, particularly in a school setting where time and cost considerations are essential. A new self-report measure (such as the CS4) that is time and cost efficient and is easy for the child to complete would be useful.

**METHOD**

**Participants**

Participants were 276 fourth-, fifth-, and sixth-grade students who attended any one of six public elementary schools in southwestern Ohio. Of the 276 students who initially re-
turned permission slips, 240 completed all three measures at both Time 1 and Time 2 of the study (analyses are based on the sample of 240 with complete data). Of the 240 students, 129 were female, with 90 children in the fourth grade, 55 in the fifth grade, and 95 in the sixth grade. Parental consent and child assent were required for participation in the study. Three schools were located in two suburban districts, and the remaining three schools were located in one urban district.

Students completed three measures on two different occasions approximately 10 to 14 days apart. Two of the measures—the CS4 and a modification of the PEI (Pekarik et al., 1976), the Peer Nomination Form (PNF)—assessed a child’s social skills. The third measure, the Depression Self-Rating Scale (DSRS; Birleson, 1981), assessed a child’s current mood state.

**Measures**

**CS4.** The CS4 is a 21-item measure in which children are asked to rate their own social behavior on a 5-point Likert-type scale (1 = never, 2 = hardly ever, 3 = sometimes, 4 = most of the time, and 5 = always). Item development for the CS4 occurred in stages. First, we identified the explicit purpose of developing the measure: The CS4 was developed with the intention to measure the extent to which the child is executing social behaviors or abilities (and not merely whether the abilities exist). We performed a literature search in PsychLit for articles concerning social skills in children, particularly focusing on social skill deficits and programs (developed since 1990) that targeted social skills training. On the basis of this search and on behavioral observations of children identified as having social skill deficits (across two clinical child facilities), we listed important domains of social skills, which included social initiating skills, adhering to social rules (e.g., politeness), and being aware of peers’ perceptions of the child. Then, we gathered measures (parent, teacher, child, and peer report) that assessed the domain of social skills. These measures included the Youth Self-Report Form (Achenbach, 1991b), the MESSY, the LSSP, and the Social Skills Rating Scale (SSRS; Gresham, Elliott, & Black, 1985), a 50-item rating scale in which teachers rate social behavior on frequency and level of importance in their classroom. We carefully examined the measures, identifying items that were common across instruments and were consistent with the purpose of identifying skills that children were not applying. From an original pool of approximately 50 items, we independently, and then conjointly, targeted 21 items. Elimination of items was based on redundancy with other items, substantial overlap with a construct other than social skills, items that focused more on affect than behavior (e.g., social anxiety), or an item that did not seem to provide useful information with regard to assessment and formulation of possible treatment for the child. Although a greater number and wider range of items could have been used on the scale, the scale is intended to be brief, with a primary focus on those critical areas directly related to the construct of social skills.

In scoring the CS4, each item is awarded 1 to 5 points, depending on the response of the child on the 5-point scale. For the seven items that are framed to measure poor skills (e.g., speaking too loudly, hitting), points awarded to a response are reverse scored. Points are added to obtain a total score, with high scores representing greater social skills. The range of possible scores is 21 to 105.

**DSRS.** The DSRS (Birleson, 1981) is an 18-item measure of depression with a 3-point response format (most of the time, sometimes, and never), with items referring to nondepressed activities (e.g., “I sleep very well”) ranging from 1 to 3, and items referring to depressed activities (e.g., “I feel like crying”) ranging from 3 to 1. These responses are totaled, so the range of possible scores is 18 to 54, with higher scores representing higher levels of depression. In the present study, the DSRS was used as a method of evaluating construct validity of the CS4 scores (i.e., children demonstrating a depressed mood should have lower scores on the CS4).
The original normative sample for the DSRS consisted of depressed and nondepressed children who ranged in age from 7 to 13 years. The 18 items of the final scale were chosen from an initial pool of 37 items on the basis of their ability to discriminate between depressed and nondepressed children. The test–retest reliability was .80, and the internal consistency reliability based on split-half was .86. The DSRS has been selected for use in other studies on the basis of its simple response format that may allow for easier comprehension for children (Friedt, 1990). Asarnow and Carlson (1986) examined the usefulness of the DSRS in a study of 82 children who ranged in age from 6 to 13 years and found that the DSRS could be used to correctly classify 77% of the children. This classification rate was replicated in another study using a larger sample of 155 children between the ages of 8 and 14 years (Birleson, Hudson, Buchanan, & Wolff, 1987). Also, high correlations between DSRS scores and Children's Depression Inventory (Kovacs, 1981) scores ($r = .81$) suggest the instrument is adequate for use in measuring childhood depression.

**PNF:** The PNF, a modification of the PEI (Pekarik et al., 1976), uses peer nominations as a method for fellow classmates to report on peers' social skills in the classroom. In the present study, the PNF was used as an additional method of evaluating the children's social skills. In other words, children who rate themselves as having good social skills should be popular among peers.

The PNF is an Item × Peer matrix, in which the names of classmates (i.e., potential nominees) are listed across the top of the page, and 11 items pertaining to social skills are listed in the left margin. When collected in the classroom, peer nominations often ask students to rate peers on a limited number of items, so that the student is able to manage the task (e.g., imagine a child considering each of 20 students for more than 15 items—this results in a possibility of 300 nominations). Examples of items on the PNF include Item 7, “Who listens well to others?” and Item 11, “Who hurts others by kicking, hitting, or name calling?” Only those students participating in the study were listed across the top. The children were instructed to place an X under the name of each child to whom the item applies and were told to nominate as many fellow classmates as they felt best fit the item. The number of Xs a child received from the other children's nominations determined his or her score on the PNF. A child received +1 for each positive nomination (i.e., an X marked for that child on a positively worded item) and −1 for each negative nomination (i.e., an X marked for that child on a negatively worded item). These numbers were then added across all PNFs and divided by the number of students from that classroom participating in the study in order to yield an average score for each child.

It is important to emphasize that the reason for gathering peer nomination scores is to assess how the children's classmates viewed each child's social skills. If a child received primarily positive nominations, the score would be high, representing good social skills. If the child received primarily negative nominations, the score would be low, indicating poor social skills. If the child received an equal number of positive and negative nominations, this would indicate that the child possessed both good and poor social skills (perhaps it was a child who was only nice to his or her close friends and then cruel to other classmates). The child's score in essence would reflect the finding that the child had neither overall good nor overall poor social skills, but rather the skills were situation or person specific. When using the average scoring method (including constructing two scores—a positive nomination score and a negative nomination score—and then combining the two), the possibility of yielding a score of 0 can occur. Although this situation did not occur in the present study, the use of this scoring method should be used with caution because information can be lost when calculating an average.

**Procedure**

The first author and a research team delivered permission slips to teachers in fourth-, fifth-, and sixth-grade classrooms in six schools (involving three school districts). The
permission slips were sent home to the students’ parents. Approximately 1 week later, permission slips were gathered; only those students who returned a signed slip from a parent or legal guardian were permitted to participate in the study. A trained research team, led by the first author, went to each of the schools and administered the three measures. The order of the social skills measures was administered randomly by class and the DSRS was always administered last, leaving two possible orders (e.g., PNF, CS^4, then DSRS; or CS^4, PNF, then DSRS). Although no order effects were anticipated, the DSRS was administered last to avoid the remote possibility that completing the depression measure would induce a sad mood state and possibly influence responding on the social skills measures. Scripts provided structured responses for anticipated questions from the children. Ten to 14 days after the initial administration the team returned to the same classrooms and readministered the three measures. At both times, the directions and the individual items of each measure were read aloud by the researcher, to ensure that children of all reading abilities were able to understand all the questions.

RESULTS

The descriptive analyses reported are based on the scores derived from the first administration of the measures (N = 240). Overall, the scores on the CS^4 ranged from 45 to 105, with a mean score of 81.36 (SD = 10.14). The scores on the DSRS ranged from 18 to 42, with a mean score of 28.69 (SD = 4.92). PNF average scores ranged from -3.0 to 4.6, with a mean score of 1.23 (SD = 1.33). Table 1 provides the means and standard deviations for the CS^4 by grade and by school district.

A one-way analysis of variance was performed to determine if there were differences in the students’ scores as a function of the district in which the child attended school. Results demonstrated that there was not an overall effect with regard to district, F(2, 237) = 2.619, p = .08, R^2 = .02. In addition, no significant differences on the CS^4 scores were found among the different grade levels, F(2, 237) = 2.95, p = .06, R^2 = .02. Because no significant differences were found in CS^4 scores across grades and school districts, analyses will be presented for the sample as a whole.

Principal Component Analysis of the CS^4

In order to explore the underlying structure of the data from the CS^4, an exploratory principal component analysis (PCA) was computed to reduce the data for the 21-item scale. Horn’s Parallel Analysis (HPA; Horn, 1965) was conducted to determine the appropriate number of components to retain for the CS^4. In order to conduct HPA, macro language in SPSS (Version 9) was used to

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined (N = 240)</td>
<td>82.05</td>
<td>10.78</td>
</tr>
<tr>
<td>Fourth grade (n = 90)</td>
<td>83.57</td>
<td>9.94</td>
</tr>
<tr>
<td>Fifth grade (n = 55)</td>
<td>80.47</td>
<td>11.56</td>
</tr>
<tr>
<td>Sixth grade (n = 95)</td>
<td>81.64</td>
<td>10.21</td>
</tr>
<tr>
<td>Suburban^* (n = 68)</td>
<td>83.34</td>
<td>9.68</td>
</tr>
<tr>
<td>Suburban^b (n = 51)</td>
<td>79.67</td>
<td>11.20</td>
</tr>
<tr>
<td>Urban (n = 121)</td>
<td>81.71</td>
<td>9.31</td>
</tr>
</tbody>
</table>

^*Suburban = District 1. ^Suburban = District 2.
create an artificial data set that had the same number of subjects \((N = 240)\) and variables \((n = 21)\) as the CS4 data set. In comparing the eigenvalues obtained through the original PCA and the eigenvalues obtained through HPA, the investigators are supposed to retain only those components that are larger than what one would observe by chance (i.e., in a random data set with variables that are uncorrelated in the population). Applying HPA to five random data sets has been recommended (Velicer, 1976). Thus, five random data sets were subjected to HPA, and each time HPA revealed that there were three components in the CS4 data set.

Upon determining the existence of three components, a PCA with an oblimin rotation with three components was conducted. Table 2 presents a list of the items included in the components that emerged from the analysis. All three components had eigenvalues greater than 1 and accounted for at least 7% of the variance in the component solution. Only those items with a pattern coefficient of an absolute value of .4 or above on a given component were considered to be included as an item on that component. Only one item (Item 1) did not meet this criterion. Component 1 was named Social Rules because items from this component seemed to generally measure a child’s adherence to social rules and politeness, (e.g., “I take turns with others”).

**TABLE 2**

<table>
<thead>
<tr>
<th>Item</th>
<th>Components</th>
<th>Communality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I look others in the face when they talk.</td>
<td>.35</td>
<td>-.16</td>
</tr>
<tr>
<td>3. I say thank you when someone does something nice for me.</td>
<td>.49</td>
<td>-.12</td>
</tr>
<tr>
<td>4. I kick or hit someone else if they make me angry.</td>
<td>-.62</td>
<td>.11</td>
</tr>
<tr>
<td>5. I am bossy.</td>
<td>-.41</td>
<td>.00</td>
</tr>
<tr>
<td>6. I take turns with others.</td>
<td>.70</td>
<td>.00</td>
</tr>
<tr>
<td>9. I listen to others when they talk.</td>
<td>.65</td>
<td>.02</td>
</tr>
<tr>
<td>10. I share games and toys with others.</td>
<td>.57</td>
<td>.20</td>
</tr>
<tr>
<td>11. I say I'm sorry when I hurt someone by accident.</td>
<td>.64</td>
<td>-.12</td>
</tr>
<tr>
<td>12. When I see others playing a game I would like to play, I ask if I can join them.</td>
<td>.44</td>
<td>.25</td>
</tr>
<tr>
<td>13. I make friends easily.</td>
<td>-.60</td>
<td>.01</td>
</tr>
<tr>
<td>14. I help others when they need help.</td>
<td>.63</td>
<td>.00</td>
</tr>
<tr>
<td>20. I ask others to play.</td>
<td>.44</td>
<td>.32</td>
</tr>
<tr>
<td>2. Others like me and have fun with me.</td>
<td>.00</td>
<td>.70</td>
</tr>
<tr>
<td>15. I walk up to others and start conversations.</td>
<td>.00</td>
<td>.80</td>
</tr>
<tr>
<td>16. Others do not like me.</td>
<td>-.16</td>
<td>-.45</td>
</tr>
<tr>
<td>18. Others ask me to play.</td>
<td>.18</td>
<td>.74</td>
</tr>
<tr>
<td>7. When I come over, others ask me to move or give them more space.</td>
<td>.01</td>
<td>-.31</td>
</tr>
<tr>
<td>8. I don't play fairly.</td>
<td>-.28</td>
<td>-.12</td>
</tr>
<tr>
<td>17. I speak or interrupt if someone else is talking.</td>
<td>.00</td>
<td>.36</td>
</tr>
<tr>
<td>19. I help others when they need help.</td>
<td>-.16</td>
<td>-.45</td>
</tr>
<tr>
<td>21. I am too loud when I talk.</td>
<td>-.38</td>
<td>-.32</td>
</tr>
</tbody>
</table>

Eigenvalue | 4.48 | 2.27 | 1.50
Variance accounted for | 21.82 | 10.79 | 7.00

Note. These results were obtained using a principal component analysis with an oblimin rotation \((N = 240)\). Bolded pattern coefficients indicate the highest coefficient for each item.
second component was named Likeability because the items from this component referred to the child's perceived popularity, or lack of popularity, with peers (e.g., “Others like me and have fun with me”). The third component was named Social-Ingenuousness because items included on this component generally seemed to assess poor recognition and/or comprehension of the finer details of social interactions; the basic rules may be there, but the application is lacking finesse. The items in this component may include the type of social skill deficits that are similar to those that children with attention-deficit hyperactive problems exhibit (e.g., “When I come over, others ask me to move or give them more space”).

Test-Retest and Internal Consistency Reliability

Pearson’s r demonstrated that test–retest reliability coefficients (over a 10–14-day period) between the scores obtained on the first and second administrations of all three measures were statistically significant: CS4, $r = .74, p < .001$; PNF, $r = .85, p < .001$; DSRS, $r = .73, p < .001$. The Cronbach’s alpha coefficient for the total scale score was .96. The items of the CS4 were also evaluated through calculation of correlations between individual items and the overall scale. Of the 21 items, all 21 correlated significantly with the CS4 overall scale, with 19 items demonstrating correlations above .30 ($p < .05$), meeting Streiner and Norman’s (1995) criterion for a “good” item–scale correlation. Items 7 ($r = .15$) and 21 ($r = .27$) fell below this criterion, although these correlations were still statistically significant.

The mean, standard deviation, and Cronbach’s alpha coefficient for each subscale (component) score follow: Social Rules ($M = 41.42, SD = 4.68, \alpha = .90$), Likeability ($M = 14.24, SD = 2.43, \alpha = .86$), and Social-Ingenuousness ($M = 12.61, SD = 3.45, \alpha = .84$).

Construct Validity

Correlations between the three measures, CS4, DSRS, and PNF, were statistically significant at both administrations. Pearson’s r demonstrated an overall statistically significant negative correlation between a child’s score on the DSRS and on the CS4 at both Time 1 ($r = -.41, p < .001$) and Time 2 ($r = -.34, p < .001$), indicating that higher self-reported social skills were associated with lower self-reported levels of depression. Overall statistically significant correlations were also found between a child’s score on the CS4 and the PNF rating at Time 1 ($r = .31, p < .001$) and Time 2 ($r = .30, p < .001$), indicating that high self-reported social skills were associated with positive peer nomination scores.

In addition, the relationships among the three components of the CS4 (Social Rules, Likeability, and Social-Ingenuousness) and the PNF and DSRS were examined. Each of the components correlated statistically significantly with the peer-rated social skills measure and the depression measure. Specifically, Component 1 (Social Rules) correlated statistically significantly with the PNF ($r = .27$) and the DSRS ($r = -.25$), Component 2 (Likeability) correlated significantly with the PNF ($r = .30$) and the DSRS ($r = -.34$), and Component 3 (Social-Ingenuousness) correlated statistically significantly with the PNF ($r = -.26$) and the DSRS ($r = .23$). All of these correlations were statistically significant at the .01 level.

DISCUSSION

The purpose of the present study was to develop a brief measure for children to self-report their use of social skills. Analyses demonstrated that the CS4 has potential to be used as a screening instrument in measuring a child’s social skills. Overall, the test–retest reliability of .74 for the CS4 scores was acceptable for use as a research instrument, but it should be higher for use in decision making. The internal consistency of .96 for the CS4 total scale score was high, suggesting that children’s responses were consistent across items.
The validity coefficients among the CS4 scores and the depression and peer-nominated social skills scores were low, accounting for only approximately 9% to 16% of the variance. Relative to the peer-nominated social skills scores, lower correlations would be expected with the depression scores, because although there is overlap between social skills and depression, they are not the same construct. However, higher associations among the peer-nominated scores and the self-reported scores of social skills would be expected, given they were intended to measure the same construct. The low correlations suggest that the CS4 and the PNF may be tapping into overlapping but different aspects of social skills. Although the present study followed a similar method of establishing construct validity as has been used for other self-report social skills measures, teacher-rated measures of social skills, such as the SSRS, have used scores on behavior problem checklists and academic performance to establish validity. Perhaps the addition of teacher behavior ratings in future studies with the CS4 would help account for a larger proportion of the variance in the scores.

A PCA was conducted in the present study that improved on previous methods used for determining the number of components underlying self-report social skills measures. Specifically, developers of the MESSY reported a PCA on the 62-items using a varimax rotation, which yielded a five-factor solution. Criterion for retaining a factor included achieving an eigenvalue greater than 1 (Kaiser criterion); pattern coefficients of .30 were retained for all the factor-based scales. Although the Kaiser criterion is widely used for the retention of factors or components, recent methodological analyses have demonstrated that this procedure is inaccurate and likely to result in retention of too many factors or components (Velicer, Eaton, & Fava, 2000; Zwick & Velicer, 1986). In fact, replication of the MESSY PCA (using the same component retention criterion) was attempted, and support for only the first two factors (Appropriate Social Skills and Inappropriate Assertiveness) was found (Spence & Little, 1990). In the present study, HPA revealed three components—Social Rules, Likeability, and Social-Ingenuousness—underlying the structure of the CS4. Two of these components, Social Rules and Social-Ingenuousness, seem similar to the two factors replicated with the MESSY: Appropriate Social Skills and Inappropriate Assertiveness. This may suggest that these two components represent important factors in social skills. The Kaiser criterion also was applied with a PCA on the LSSP, which indicated an eight-component solution: Social Anxiety, Strangers, Temper Control, Social Discomfort, Conflict Situations, Parents, Friendship, and Opposite Sex. Only one of the LSSP components (Friendship) appears similar to a CS4 component (Likability), but as discussed earlier, it seems that many of the LSSP items are assessing situations more related to social anxiety than to social skills.

The present study provides a foundation for the development of the CS4 for use as a screening instrument; however, a few limitations are salient. Time constraints imposed by the administration of the participating schools prevented the use of a more comprehensive battery of measures. Construct validity should then be further established by correlating the CS4 scores with the scores of other well-established measures of social skills, such as the SSRS. Also, some of the wording of the CS4 items may be difficult for children to follow if instructors (to clarify meaning) are not available. Specifically, Items 8 and 16 are worded in a negative fashion and may be difficult for children to know whether to rate high or low. Future studies should investigate the possibility of rewording these items to keep potential confusion to a minimum.

In conclusion, the CS4 is a tool that may aid in the early identification of the lack of utilization of social skills in three areas identified in this study (i.e., social rules, likeability, and social-ingenuousness). Schools have begun to focus on screening groups of children for problems and at-risk behaviors. Thus, in following the multiple-gate procedure described earlier, brief measures are necessary in the early stages of assessment to evaluate children in the most time- and cost-efficient manner. Ultimately, a good screening instrument to be used in any school setting should be complete and quick to administer and should cause the least disruption in the classroom. The CS4 can be used as a research instrument in this context but requires further investigation to determine whether its scores achieve the necessary reliability and validity conditions for clinical decision making.
REFERENCES


