International Journal of Contemporary Educational Research
Volume 12, Number 1, March 2025, Page 39-55 Article Type: Research Article

Developing the "Social Skills Assessment Scale–Teacher Form" for Preschool Children: A Validity and Reliability Study

Pınar Aksoy[1] | | aksoypnr@gmail.com
Tokat Gaziosmanpaşa University, Faculty of Education, Department of Preschool Education, Tokat, Türkiye
Gülen Baran | | barangln@gmail.com
Ankara University, Faculty of Health Sciences, Department of Child Development, Ankara, Türkiye

Abstract
This study aimed to develop the "Social Skills Assessment Scale–Teacher Form" (SOSAS–TF) for preschool children. Data were collected from three distinct groups of participants. The first dataset, obtained from 254 preschool children, was collected during a *preliminary trial*. The second dataset, gathered from 896 preschool children, was used for the *trial practice*, while the third dataset, consisting of 263 preschool children, was used to assess *test-retest reliability*. Within this scope, the Kaiser-Meyer-Olkin (KMO) value was determined to be .96, and Bartlett's test yielded a significant chi-square value, indicating the suitability of the data for factor analysis. The findings of the study revealed that the scale, consisting of 44 items, included five dimensions: "communication," "adaptation,'' "self-control,'' "prosocial behaviors,'' and "assertiveness.'' The factor loadings of all items ranged from .46 to .74, while the common factor variances varied between .42 and .69. These five factors collectively explained 54.57% of the total variance. The Cronbach's alpha reliability coefficients (α) for the dimensions of "communication," "adaptation," "self-control," and "prosocial behaviors" were .89, .90, .86, and .87, respectively. The total reliability was strong (α = .94), and the "assertiveness" dimension (α = .76) was found to be significantly reliable. Regarding the scale's test-retest reliability, the correlation coefficients for the dimensions of "communication" (r = .82), "adaptation" (r = .86), "self-control" (r = .75), "prosocial behaviors" (r = .86), and "assertiveness" (r = .63), as well as the total score (r = .89), were positive and highly significant (p < .001). In conclusion, a valid and reliable assessment instrument was developed, contributing to the literature on social skills assessment in preschool children.

Keywords: Preschool children, Social skills, Social Skills Assessment Scale–Teacher Form, Scale development, Validity and reliability

Citation

Aksoy, P. & Baran, G. (2025). Developing the "Social Skills Assessment Scale─Teacher Form" for preschool children: A validity and reliability study. *International Journal of Contemporary Educational Research*, *12*(1), 39-55. https://doi.org/10.52380/ijcer.2025.12.1.390

Received	14.03.2024
Accepted	19.03.2025
Publication	28.03.2025
Peer-Review	Double anonymized - Double Blind
Ethical Statement	* This article is derived from the corresponding author's doctoral dissertation, entitled "Examination of the Effect of Story Telling Based and Play Based Social Skills Training on the Social Skills of Children Attending to Kindergarten," supervised by the second author (Ph.D. Dissertation, Ankara University, Ankara, Turkey).
Plagiarism Checks	Yes - iThenticate
Conflicts of Interest	The author(s) has no conflict of interest to declare.
Complaints	editor@ijcer.net
Grant Support	The author(s) acknowledge that they received no external funding in support of this research.
Copyright & License	Authors publishing with the journal retain the copyright to their work licensed under the **CC BY-NC 4.0**.

[1] Corresponding Author

Introduction

Social skills are considered learned behaviors that are socially accepted and allow individuals to respond positively to their interaction with others, and help them avoid negative reactions (Gresham & Elliot, 1984). Social skills, in a way, are certain behaviors that are required for successfully achieving social duties (McFall, 1982) and serve as guiding principles for individuals to quickly reach their goals, take advantage of opportunities, and adapt to their environment easily (Gresham, 2002; Gresham & Elliott, 1987). At this point, social skills enable individuals to achieve acceptance in social life and encourage them to be in harmony with their environment. Similarly, as individuals' mental health is rooted in establishing healthy relationships with others, social skills are highly important for protecting individuals' mental well-being. In the case of inadequacies in social skills, individuals' satisfaction with close relationships tends to be lower, and they are more likely to experience increased clinical problems such as depression, loneliness, and social anxiety (Arnold et al., 2012; Campbell et al., 2010; Ceylan, 2009; Clayton et al., 2021; Dong et al., 2023; Erkul & Sonmez, 2020; Gresham, 2001; Gresham et al., 2006; Jones et al., 1982; Kalutskaya et al., 2015; Leary & Kowalski, 1995; Malecki & Elliot, 2002; Mulvey et al., 2020; Segrin, 2000; Segrin & Flora, 2006; Sharma et al., 2016; Welsh et al., 2001; Whirter & Voltan Acar 1998; Zsolnai, 2002). Inadequacies in social skills hinder the development of interpersonal communication and prevent individuals from establishing effective communication with people around them in a social environment, thus reducing their participation in social settings. These deficiencies, more importantly, prevent individuals from forming meaningful relationships with their environment, ultimately leading to regression in life and social isolation (Bellini, 2008; Conger & Keane, 1981; Erwin, 1994; Gresham, 2016; Gresham & Elliott, 1987; Kalutskaya et al., 2015; Lodder et al., 2016; Quinn et al., 1999; Strain et al., 1984). Children who struggle with social interactions often experience difficulties in socio-emotional development, which may lead to lower academic achievement and heightened vulnerability to emotional challenges (Kılıç et al., 223; Offer-Boljahn et al., 2022; Parker & Asher, 1987; Rademacher et al., 2020; Schneider & Byrne, 1985; Zöggeler-Burkhardt et al., 2023). Liberman et al. (1989) identified several factors contributing to inadequacies in individuals' social skills, as follows. (i) A lack of appropriate role models, preventing individuals from learning effective social interactions; (ii) conditions such as social anxiety, alcohol addiction, and schizophrenia, impairing the development of social skills; (iii) traumatic events and stressful situations, making it difficult for individuals to demonstrate their social skills effectively; and (iv) significant life changes, such as imprisonment, homelessness, job loss, starting preschool, and relocating to a new environment.

When an inadequacy in social skills is detected, targeted interventions are necessary to address the deficiency. At this point, employing effective strategies and intervening during the early stages of life is crucial. The primary approach to addressing social skills inadequacies is identifying their underlying causes. If the inadequacy results from a lack of appropriate role models, it may be necessary to create environments that provide children with examples of appropriate behavior or to organize activities that allow them to observe such behaviors. Additionally, if the deficiency stems from individual characteristics that hinder the acquisition or demonstration of social skills, solutions should focus on addressing the specific behaviors or skills contributing to the problem. Gaining insights into these challenges requires a comprehensive and systematic evaluation (Cooke & Apolloni, 1976; Dong et al., 2023; Günindi, 2023; Gresham, 2001; Gresham et al., 2006; Meuwissen, 2022; Takahashi et al., 2015).

As social skills are highly complex, various techniques, such as formal and informal methods, are required to evaluate them. Standard measurement instruments are used for formal evaluations, while observation, interviews, sociometry, self-evaluation, and behavior-rating scales are used for informal evaluations (Cooke & Apolloni, 1976; Kutlu & Kaya, 2005). Sociometry is a useful method for evaluating social skills, but it has some limitations in how it can be used comprehensively in line with the objectives (Cornish & Ross, 2004; Merrell, 2001). This explains why sociometric techniques, although commonly used, may result in less peer acceptance for children. Results from sociometric evaluations may vary depending on the classroom population and do not provide adults with information on which social skills of the children should be supported. Therefore, it is recommended to use sociometry in conjunction with other measurement instruments (Foster et al., 1993; Kavale et al., 1988; Warnes et al., 2005). The self-evaluation technique involves individuals marking whether they have performed a behavior or not, or to what extent they have performed it, using rating scales or checklists. In this sense, the individual makes an evaluation based on their judgment. Using the self-evaluation technique alone to determine inadequacies in social skills is not considered sufficient, and it is recommended to use it in combination with additional information obtained through direct observation. Nonetheless, it is not possible to apply the self-evaluation technique to preschool children, as the concrete thinking and self-centered mentality of preschool children make it difficult for

them to make objective evaluations about themselves (Bacanlı, 2008; Connolly & Doyle, 1981; Deutsch, 1974; Elliott, & Busse, 1991; Gülay & Akman, 2009; Merrell, 2001; Merrell & Gimpel, 2014; Warnes et al., 2005). Although natural observation has some advantages, such as not involving intervention and being specific and objective, it also has disadvantages, such as susceptibility to bias and lack of predictive validity. Based on this, it is not recommended to use a single instrument solely for the evaluation of social skills. Moreover, certain target behaviors can only be observed in particular environments, and thus, it is not sufficient to evaluate less exhibited behaviors with a single observation. Therefore, it is also important to use broad screening inventories and social skill rating scales to make a comprehensive evaluation of social skills in children (Boisjoli & Matson, 2009; Elliott, & Busse, 1991; Gresham, 1981; Sheridan, & Walker, 1999).

Behavior-rating scales provide several advantages for evaluating children's social skills. Merrell (1999) suggested that the advantages of behavior-rating scales can be categorized under six dimensions. First, behavior-rating scales require less professional time and training activities to utilize the evaluation system compared to direct behavioral observation. Second, while data collected through behavior-rating scales might be of low frequency, it is possible to obtain data on important behaviors that might not be revealed during a limited number of direct observations. Third, an objective evaluation method is provided with behavior-rating scales, and this method yields more reliable data than expressive techniques or unstructured interviews. Fourth, they are useful for evaluating individuals who have difficulty providing information about themselves, including children with low verbal skills and/or young individuals who tend not to be cooperative. Fifth, behavior-rating scales can be used in children's schools or homes, which may be considered the natural environment for children. Sixth, children's parents or teachers are expected to have strong observations of and judgments about children, as they are highly familiar with them, and these observations and judgments are used in behavior-rating scales. In this respect, it is concluded that behavior-rating scales are effective and important instruments for evaluating children's social skills (Boisjoli & Matson, 2009; Hosp et al., 2003; Merrell & Gimpel, 2014). Such cases indicate that evaluating social skills through rating scales is more efficient. As highlighted, developing a reliable scale to assess preschool children's social skills would be valuable.

The evaluation of children's social skills by individuals who are familiar with these children is an ideal application for behavior-rating scales. In this context, teachers are considered one of the most reliable and valuable sources of information regarding children's behaviors. This is because teachers systematically observe and monitor children in various situations, and they are the individuals who interact with children during play and other activities. Additionally, teachers are expected to possess both theoretical and practical knowledge of child development, which enables them to assess children's behaviors and skills realistically. Furthermore, teachers are trained in supervision techniques and are knowledgeable about behaviors that deviate from the norm, thus ensuring the reliability of their evaluations (Çağlar, 1981; Connolly & Doyle, 1981; Wilson & Bullock, 1989). Nevertheless, it is crucial to recognize that social skills are behaviors exhibited in social environments and society. Teachers can observe children's behaviors in the classroom and their interactions with peers, which increases the validity of the results. Therefore, during the development of the scale in this study, validity and reliability analyses were conducted based on data provided by teachers. Similar measurement instruments are used in studies evaluating the social skills of preschool children, and currently, there are a limited number of measuring instruments available in the literature (Avcıoğlu, 2007; Gresham & Elliott, 1990; Kapıkıran et al., 2006; Merrell, 1996). It is noteworthy that these measurement instruments were developed in previous years. While social skills remain essential, their assessment should be conducted using updated processes. Each contribution to the literature on social skills evaluation represents a significant step in advancing the field. In this regard, the measurement instrument developed in this study to assess preschool children's social skills is expected to serve as a valuable resource for future research and provide researchers with a new perspective. Accordingly, this study aimed to develop the "Social Skills Assessment Scale–Teacher Form" (SOSAS–TF) for preschool children. To achieve this, validity and reliability studies were conducted with Turkish preschool children.

For the scope of this study, existing instruments for assessing social skills were comprehensively analyzed, with the aim of developing a more contemporary and effective instrument based on the insights and findings obtained from these instruments. The existing instruments generally measure social skill levels from a broad perspective and are limited to specific observations within the educational environment. These limitations pose challenges in evaluating social skills in a more detailed and specific manner. In this context, the present study is designed to develop a psychometrically reliable instrument that better meets the needs of the education system and teacher evaluation processes, while also considering cultural sensitivities. During the scale development process, no initial limitations were imposed on the number of items. Instead, to ensure validity, items that allowed social skills to be measured in a more comprehensive, directly observable, and context-specific manner were prioritized. At this stage, adaptation to the unique characteristics of the Turkish preschool education system was prioritized, with the ultimate goal of designing a scale that would enable teachers to assess social skills in a more valid and

comprehensive way. This scale, developed in response to current needs and increasing demands in the field of education, is expected to be a practical instrument that not only determines children's social skill levels but also enables teachers to monitor and support social skills development more effectively.

Method

This study focused on developing a measuring instrument to evaluate preschool children's social skills, based on the results of exploratory factor analysis. The operations performed during the process of developing the SOSAS–TF scale are illustrated in Figure 1.

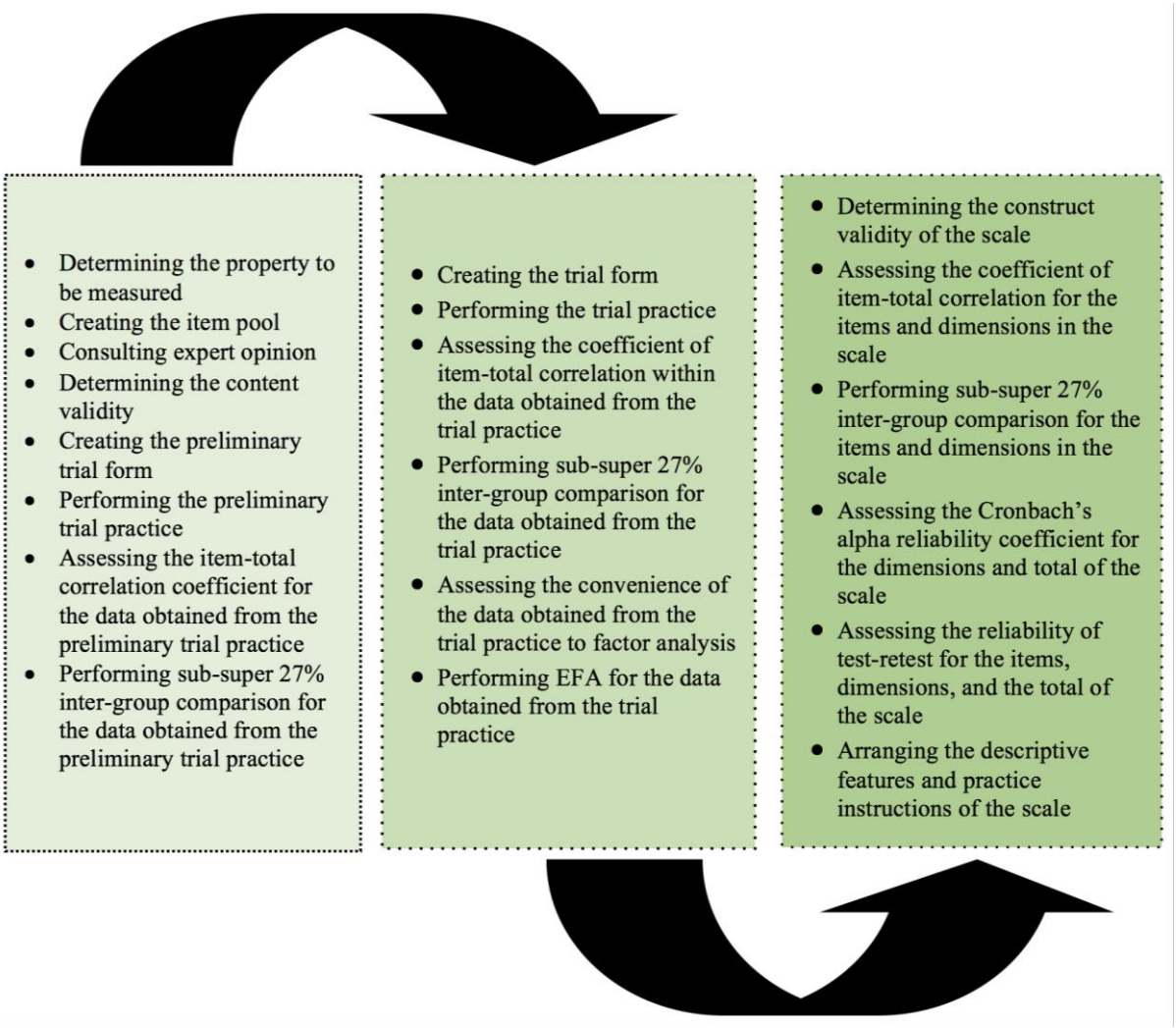

Figure 1. Operations performed during the process of developing the SOSAS-TF scale

Participants

The study was conducted with three groups of participants. In the first group, a **preliminary trial** was performed. For the preliminary trial of the items, upon determining their content validity, eight preschool teachers working at five schools in the city center of Tokat, Turkey, were asked to view the form. The teachers were asked to indicate the items they had difficulty in understanding, or they found incomprehensible. As no item was reported as incomprehensible/inappropriate, implementation of the preliminary trial was continued. Accordingly, the researcher distributed the preliminary trial form consisting of 66 items, to the teachers through one-to-one interviews and explained the researcher's purpose and the implementation instructions of the form. Accordingly, a group of preschool teachers working at 20 schools was selected among the teachers working at preschools and

kindergarten classes located in a city center in the Black Sea Region in Turkey. They were asked to evaluate the skills given in the form according to children's frequency of exhibiting the skills in their classrooms. Ten forms were given to each of the 29 preschool teachers who volunteered to participate in the study to fill in, considering the children in their classrooms. In this context, any form that was filled incorrectly was not included in the implementation data. Thus, preliminary trial form data of 254 children were collected. Among these children, 54.3% (n= 138) were boys and 45.7% (n= 116) were girls, with a mean age of 66.4 months (SD = 3.5).

The **trial practice** was performed in the second group. In factor analysis studies, it is recommended that the number of participants should be approximately five or ten times greater than the number of the items in the scale to be developed to obtain eligible results (Tavşancıl, 2005). Comrey & Lee (1992) defined the adequacy of a sample size, which will be subjected to factor analysis, as "very bad" for 50 participants, "bad" for 100 participants, "medium" for 200 participants, "good" for 300 participants, "very good" for 500 participants, and "perfect" for 1000 and more participants. According to this, it was aimed to include at least 10 times more children in the study group than the 63 items in the scale for the trial form. In this context, 92 preschool teachers working at 36 schools were selected through a random sampling method among teachers in the city center of Tokat. These teachers represented schools of different socioeconomic levels. The forms of 896 children who were evaluated correctly by these teachers, were obtained. Accordingly, the number of the forms filled out by the teachers for the children in their classrooms varies between 2 and 12, and 53.8% (n= 482) of the children, who were evaluated within the scope of the trial form, were boys; while 46.2% (n= 414) of them were girls. The mean age of this children was 65.9 months (SD = 3.7).

Test-retest reliability was performed in the third group. To determine the external reliability of the scale, the classrooms of the teachers who had previously filled out the form within the scope of the trial implementation were revisited four weeks later, and they were asked to fill out the form once more for the children who had previously been subjected to the form. In this context, the teachers were given forms for 300 children, and they filled out the forms for 263 children completely. 55.1% (n= 145) of the children who were involved in the process for test-retest reliability were girls, while 44.9% (n= 118) of them were boys, with a mean age of 66.1 months (SD = 3.4).

Process

The study was conducted after obtaining a letter of approval from the Turkish Ministry of National Education [MoNE] and consent forms from parents. During the development of the SOSAS–TF scale, a literature survey was first conducted, and then a search was made on studies regarding this subject and existing assessment instruments in the national and international literature. In this context, social skills evaluation scales for preschool children developed in various countries (Gresham, & Elliott, 1990; Merrell, 1996) and adapted to Turkish (Kamaraj, 2004; Koçyiğit, & Kayılı, 2008), and that developed in Turkey (Avcıoğlu, 2007; Gülay, 2004; Kapıkıran et al., 2006), were reviewed. At this point, the learning outcomes and indicators in the preschool program of the MoNE (2013) were addressed. Accordingly, 106 items in the scope of social skills were created. These items were examined based on several criteria: whether they explain a single characteristic/skill, their expression in a simple and understandable manner, observability in an environment, suitability for preschool children, evaluability by teachers, relevance to the study's purpose, and the contextual skills they encompass. A theoretical and conceptual framework was used to develop scale items as a data collection instrument and to create a trial form. These items were organized according to the social learning approach (Bandura, 1972; Rotter, 1982), which integrates common elements from multiple theories (cognitive and behavioral theories) and is centered on children's learning through observation and modeling.

Seventy-four candidate scale items were created by eliminating and correcting some items. In order to balance the validation trends of these items in the draft form during the evaluation phase, and avoid manipulations, 38 of them were defined as positive, and 36 were defined as negative (Tavşancıl, 2005). At this point, the evaluation criteria of the scale items in the literature are reviewed and it was decided to evaluate the items in the scale as "always," "often," "sometimes," "rarely," and "never" based on the frequency of each child to exhibit them. In addition to this, "always" was defined as 5 points, "often" as 4 points, "sometimes" as 3 points, "rarely" as 2 points, and "never" as 1 point for the items expressed positively. Reverse scoring (recode) was decided to be applied for the items that were expressed negatively, and "always" was defined as 1 point, "often" as 2 points, "sometimes" as 3 points, "rarely" as 4 points, and "never" as 5 points. A section for implementation instructions was included in the introduction of the form, including the scale items, providing explanations on the aim and the response manner of the scale.

The content validity of the 74 candidate scale items created in the study was analyzed. Content validity reveals whether the items in the scale are sufficient to measure social skills in terms of quality and quantity. Expert opinion is considered one of the logical methods of determining the content validity (Büyüköztürk, 2014). Therefore, expert opinion was obtained to test the content validity of the items. In this context, a form was given to the experts in which they report to what extent they found it appropriate to include the items in the preliminary trial form into the scale. The experts were asked to evaluate each item in the trial form using a rating coding system that included categories such as "suitable," "partly suitable," and "not suitable". They were also asked to provide a "comment" section for each item in the form, allowing them to add their opinions and recommendations about the items. They were asked to evaluate the items based on their suitability for preschool children, alignment with the scope of social skills, clarity of expression, comprehensibility, utility for teachers/parents to use, and coverage of various dimensions of social skills while determining the content validity of the items that were created in line with expert opinion.

In this direction, the opinions of 11 experts were obtained to ensure content validity, comprising five field experts in *preschool education*, four field experts in *child development and education*, and two field experts from *psychological counseling and guidance department* who conducted studies on social skills during the preschool period. Additionally, two Turkish language education field experts were consulted to assess whether the items adhered to correct grammar rules and were comprehensible. Furthermore, three field experts in measurement and evaluation analyzed the items in terms of the way of expression and their procedural convenience. Accordingly, the opinions were obtained from a total of 16 experts. Twelve of these experts had doctoral degrees, three had master's degrees, and one had a bachelor's degree. Meanwhile, fourteen of them were academics, and two were preschool teachers.

After the expert opinions were obtained, all feedback in the expert forms was combined in a single form, and the number of experts approving the potential choices of each item was determined. During this process, according to the expert opinions, items with 90% and above approval were completely selected. Items with 70-80% approval were rearranged according to the recommendations, and items with lower approval rates were removed from the scale (Büyüköztürk, 2014). In this scope, six items of the draft scale were excluded, four items were rearranged, and two items were combined using a common expression that was present in both to create a single item. According to the expert opinions, no new item was included as no item was recommended to be added. Thus, a preliminary trial form consisting of 66 items among 74 candidate scale items was created.

Data (of 254 children) on the preliminary trial form consisting of 66 items were collected in the study. As a result of the Pearson correlation analysis conducted to determine the item-total correlation, it was found that item 32 (r= -.017, p= .784), item 46 (r= .085, p= .179), and item 50 (r= .193, p= .001) had lower item-total correlation. It was determined that the item-total correlation of the remaining 63 items apart from these items (items 32, 46, and 50) was greater than .30. Items with an item-total correlation of .30 and higher distinguish individuals very well; items with an item-total correlation between .20 and .30 might be included in the trial, in case it is deemed mandatory, or the item should be corrected, and items with an item-total correlation lower than .20 should not be included in the trial.

The positive and high item-total correlation indicates that the items illustrate similar behaviors and internal consistency of the trial is high (Büyüköztürk, 2014). According to this, it was determined that items 23, 46, and 50 should be excluded from the scale, as they would have a negative effect on the internal consistency of the trial. Item distinguishing capability was also considered to decide the item exclusion process. Sub-super 27% group comparison was performed to determine the distinctiveness validity of the SOSAS–TF scale.

A t-test for independent groups was performed to determine whether the difference between the item scores of the subjects in the sub (n= 69) and super (n= 69) groups for each item of the test scale was significant. Similarly, the findings revealed that the total scores of items 32 (t= .333, p= .739) and 46 (t= -.603, p= .547) did not differ significantly among the sub and super groups subjects. Accordingly, it was decided to exclude the items 32, 46, and 50 from the scale. At this point, a high and significant difference was found in the other items in the trial form (p= <.001). The differences observed in the desired direction between the groups were found to be significant, and this is considered an indicator of the internal consistency of the test (Büyüköztürk, 2014). In this context, the 66-item preliminary trial form was converted into a 63-item trial form.

Data (of 896 children) on the preliminary trial form consisting of 63 items created in the study were collected at this stage. The item-total correlation coefficients and distinctiveness capabilities of the scale items were considered

to determine whether the items in the dataset obtained through the trial form illustrated similar behaviors and to keep the items with high internal consistency in the scale.

The findings of the Pearson correlation analysis that was performed to determine the item-total correlation revealed that items 29, 33, and 34 were found to have values lower than .30, and were excluded from the scale. Thus, they were not used in the analysis. The findings of the *t*-test analysis for independent groups that was conducted between 27% sub (n= 242) and super (n= 242) groups to determine the capability of item distinctiveness revealed that items 35, 37, 66, and 68 did not have significant differences and were excluded from the scale before the analysis. Consequently, 56 items were kept in the scale. At this stage, the validity and reliability studies were conducted on the scale.

Data Analysis

The data collected through the study were transferred to the Statistical Package for the Social Sciences (SPSS version 20.0) and made ready for the validity and reliability analysis. The following steps were followed to analyze the data at hand:

In this context, the item-total correlation and the capability of item distinctiveness were considered to demonstrate the item's validity. The item-total correlation was calculated using the Pearson product-moment correlation coefficient, and the capability of item distinctiveness was evaluated by comparing the top and bottom 27% subgroups through an independent groups *t*-test. The values of the Kaiser-Meyer-Olkin (KMO) test for sampling adequacy and Bartlett's test of sphericity were analyzed to determine whether the data at hand were suitable for factor analysis. An exploratory factor analysis (EFA) was conducted to determine the scale's construct validity. The contribution of each factor to the total variance (eigenvalues), their percentages to variance contribution, and their cumulative percentages to variance contribution were reviewed. Cronbach's alpha coefficient of reliability was considered to determine the reliability of the total scale's dimensions.The relationship between the two sets of scores that were obtained at different times was analyzed through thee Pearson product-moment correlation coefficient to determine the test-retest reliability of the scale.

Findings and Comments

This section contains the findings of the validity and reliability of the SOSAS–TF scale for preschool children. In this context, the findings of studies based on validity and reliability analyses are presented and interpreted.

Validity Studies

The success of the measurement instrument in predicting children's social skills depends highly on its validity and reliability. Validity is a concept that tries to measure to what extent the measurement instrument accurately captures the desired characteristic of the individual. One of the characteristics that should be taken into account while determining the validity of the measurement instrument is construct validity. The capability of the measurement instrument to capture an abstract concept accurately in the context of the desired behavior represents the construct validity. In this study, a large number of measurable and observable items were created to measure the social skills of the individuals. The results reveal to what extent these items measure the characteristics of social skills are related to the construct validity. Factor analysis, cluster analysis, internal consistency, and hypothesis testing techniques can be used to determine the construct validity (Çokluk et al., 2014; Kline, 2014; McDonald, 2014).

Factor analysis aims to express measurement with a small number of factors, aggregating the variables that measure the same structure or qualification. Factor analysis is the process of revealing new variables called factorization or common factor, or in the process of obtaining functional definitions of the concepts, using factor load values of the items. There are two types of factor analysis: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA).

Exploratory factor analysis contains an activity to find a factor based on the relations among variables. Thus, EFA is a technique to understand the existing construct. However, CFA contains an activity to test a hypothesis or theory about the relationship among the study variables. Exploratory factor analysis (EFA) is used to determine the construct validity of the scale that is developed as a data collection instrument in social sciences (Büyüköztürk, 2014; Erkuş, 2012; Finch, 2019). Under these circumstances, it was first evaluated whether the obtained dataset was suitable for factor analysis and the data had a multivariate normal distribution. Before starting factor analysis,

the Kaiser-Meyer-Olkin (KMO) test for sampling adequacy and Bartlett's test of sphericity values were reviewed to determine whether the data at hand were suitable for factor analysis. As a result of the analysis, it was determined that the KMO value, which analyzed the suitability of the group for the factor analysis in terms of the size of the sample, was .96 which is the threshold for adequate sample size, had a value between 0 and 1. When the KMO value is lower than .50, it means that the dataset is not suitable for factor analysis (Cerny & Kaiser, 1977).

When the KMO value is lower than .50 in terms of its suitability for factor analysis, it means it is an "unacceptable" situation. The value .50 represents "weak," .60 represents "medium," .70 represents "good," .80 represents "very good," and .90 represents "perfect" (Sharma, 1995). Hence, the closer the KMO value to 1.00, the higher the suitability of the data for analysis (Tabachnick & Fidell, 2013). The KMO value of the study was found to be .96 which indicates that the sample size of the study is perfectly suitable for factor analysis. Similarly, according to the result of the Bartlett test, the approx. chi-square (X_2) value was significant at the significance level of .01 [$X^2(1711) = 27060.610$; $p < .001$]. This result showed that the data originated from the multivariate normal distribution, and another premise for the factor analysis was met (Çokluk et al., 2014). The data are considered suitable for factor analysis when the KMO value is higher than .60, and the result of the Bartlett test is found to be significant. These results suggest that the data at hand could be factorized (Büyüköztürk, 2014; Pett et al., 2003; Tabachnick, & Fidell, 2013). Therefore, it was decided that the dataset was suitable for factor analysis. Accordingly, 56 items were subjected to EFA to determine the construct validity of the scale, and the results of the total explained variance were analyzed. At this point, each factor's contribution to the total variance (eigenvalues), percentages to the variance contribution, and cumulative percentages to the variance contribution were examined. There was no limitation to the number of the factors in EFA and the varimax orthogonal rotation technique was used to find a more readable or suitable factor construct without damaging the explained variance.

As a result of the varimax rotation performed to determine the number of factors that constituted the scale and what these dimensions were, it was found that there were nine factors with eigenvalues higher than 1.00, and the nine factors explained 57.01% of the variance. However, it was remarkable that the contribution to the variance from the sixth component was not important. At this stage, the scree plot related to the factor eigenvalues was analyzed, and it was found that there was no change in the graphic curve from the sixth component. This suggests that the items in the scale could be categorized under five factors.

The load values of the items in the factor they are involved, which are generated as a result of EFA, are considered a good score for selection when they have values equal to or more than .40, according to Tavşancıl (2005) and .45 or more according to Büyüköztürk (2014). At this point, it is stated that the presence of a cluster constituted by items, which yield a high-level relationship with the factor, expresses that the items together measure a construct. Considering the factor load value was .45, within the scope of the study, the analysis was continued. Therefore, five items that were below the threshold (items 9, 26, 27, 50, and 65) were excluded from the scale. Afterward, it was observed that five items with acceptable load values in more than one dimension overlapped (items 3, 5, 28, 32, and 59). This shows that these items were related to more than one factor and cannot be collected under only one factor. It is recommended that the difference between the highest load value of an item in the factor and the second highest load value be at least .10 (Büyüköztürk, 2014). Therefore, considering the situation where the difference among the factor loadings was at the lowest level, these items were excluded from the scale items one by one. At this point, two items that were not adequate to create a factor in terms of quantity and that were collected under the same factor were excluded from the scale (items 10 and 66), then the analysis was repeated. Hence, the common variance loads of the remaining items in the scale were examined, and it was determined that 44 items that could be explained by five factors, would remain in the scale.

At this point, the explained variance and the contribution to the total variance by each factor were analyzed. According to Kalaycı (2014), in case the contribution of each additional factor to the explanation of the total variance decreases below 5%, it indicates that the maximum number of factors is reached. Hence, it was determined that the first factor explained 15.54%, the first two factors explained 27.01% together, three factors explained 37.61%, four factors explained 47.04%, and five factors explained 54.57% of the total variance. It was then concluded that the contribution of the first five components to the variance was at an adequate ratio.

The variance rates varying between 40% and 60% in the multifactor patterns are accepted to be adequate (Scherer et al., 1988; Tavşancıl, 2005). In this framework, it appears that the contribution of the five-factor construct to the total variance is adequate. Here, it was observed that the first factor with the eigenvalue of 6.83 explained 15.54%, the second factor with the eigenvalue of 5.04 explained 11.47%, the third factor with the eigenvalue of 4.66 explained 10.59%, the fourth factor with the eigenvalue of 4.15 explained 9.43%, and the fifth factor with the

eigenvalue of 3.31 explained 7.52% of the total variance. Table 1 shows the EFA results of the five-factor construct.

Table 1. Results of the EFA related to the SOSAS-TF scale and the result of the *t*-test between 27% of sub-super groups

Draft scale	Final scale	The first factor	The second factor	The third factor	The fourth factor	The fifth factor	Common variance	Result of t-test between 27% of sub-super groups
m8	m1	.719					.658	19.81***
m20	m2	.676					.636	23.98***
m13	m3	.675					.572	21.75***
m4	m4	.644					.557	21.86***
m31	m5	.640					.565	22.49***
m12	m6	.630					.574	18.83***
m56	m7	.624					.517	19.66***
m2	m8	.623					.501	18.72***
m22	m9	.604					.577	24.88***
m38	m10	.583					.512	15.27***
m1	m11	.528					.501	19.31***
m30	m12	.487					.482	15.30***
m49	m13		.725				.666	18.98***
m48	m14		.724				.694	17.03***
m43	m15		.693				.662	14.69***
m42	m16		.657				.639	18.89***
m61	m17		.618				.550	13.65***
m55	m18		.599				.610	20.53***
m41	m19		.587				.504	12.43***
m52	m20		.550				.625	25.86***
m47	m21		.522				.536	18.70***
m11	m22			.742			.621	16.97***
m54	m23			.719			.562	12.58***
m16	m24			.641			.519	15.50***
m37	m25			.641			.503	9.22***
m45	m26			.628			.529	13.04***
m14	m27			.596			.484	15.70***
m19	m28			.564			.516	21.94***
m63	m29			.532			.476	12.71***
m23	m30			.525			.597	21.53***
m25	m31				.628		.564	19.06***
m58	m32				.621		.579	19.53***
m17	m33				.603		.621	21.02***
m36	m34				.583		.597	19.63***
m53	m35				.580		.502	15.60***
m51	m36				.571		.546	19.12***
m18	m37				.557		.553	21.31***
m15	m38				.515		.482	18.84***
m21	m39				.462		.424	17.45***
m69	m40					.733	.649	14.80***
m24	m41					.661	.563	16.76***
m64	m42					.613	.468	8.51***
m6	m43					.586	.527	10.73***
m7	m44					.550	.514	15.45***

*** p <.001

Table 1 shows that the factor loadings in the first factor vary between .49 and .72, the factor loadings in the second factor vary between .52 and .73, the factor loadings in the third factor vary between .53 and .74, the factor loadings in the fourth factor vary.46 and .63, and the factor loadings in the fifth factor vary between .55 and .73. Thus, it was observed that the factor loadings of all of the items in the scale varied between .46 and .74, and the common factor variances, which represented the capability of each item to explain the total variance related to the scale

together with the other items, varied between .42 and .69. According to this, it is remarkable that the factor loadings of all of the items remained in the scale, and they all had .45, which is recommended as a good threshold to measure the construct; and the common factor variances were higher than .30, which is accepted as an adequate value (Büyüköztürk, 2014; Kline, 2014). These findings show that the items on the scale would suggest the phenomenon of social skills well.

At this stage, the content of the items collected under the same factor and the factor construct represented by them were examined, which led to the opinion that the items under a common factor were the ones that measured the characteristics, which were in a similar construct. At this point, naming the resulting five factors was emphasized. For this, as suggested by Kalaycı (2014), the items of the same factor and the items that had major weight under a factor were examined. According to this, the first dimension was named as "communication" dimension as it included items such as "give right answers to the questions about other's feelings," "expresses their feelings clearly in the face of an event or situation," "arranges their gestures and facial expressions for the situation by an appropriate way while talking," and "makes eye contact with the listener while talking."

The second dimension was named as "adaptation" dimension as it includes items such as "follow the given instructions," "follow the rules in his/her environment," and "wait for his/her turn in the situations that require moving in turn." In the meantime, the third dimension was named as "self-control" dimension as it contained items such as "show aggressive behaviors when s/he gets angry," "respond in the same way when others pushed/hit," and "damage the objects in his/her class/room." The fourth dimension was named as "prosocial behaviors" as it contained items such as "appreciates the success of others," "express his/her love by hugging or verbally expressing his/her love," and "voluntarily help someone whom s/he feels needs help." Additionally, the fifth dimension was named as "assertiveness" as it contained items such as "abstain from joining other children's play,'' "has difficulty in making new friends,'' and "is shy and withdrawn.'' After the researcher named the factors, the opinions of five experts, three in the field of preschool education and two in the field of child development and education, were obtained to determine the appropriateness of the factor names. All experts found the factor names appropriate, and, thus, the names were accepted as the dimension names in the scale.

Reliability Studies

The consistency of the items in the scale with each other and the extent to which the scale reflects the addressed skills are related to the reliability of the scale. Reliability, in one respect, refers to the extent to which the measurement instrument accurately captures the desired characteristics. At this point, it is recommended that an item prepared to form the scale should be examined in terms of its capability to be consistent and stable in itself and cause desired reactions to be collected. The *t*-test results about the difference between the item averages of the sub 27% and super 27% groups, which were established according to the coefficient of item-total correlation and the total scores of the measurement instrument, was taken into consideration while determining the measuring capacity related to an item, and the statistical significance was considered the criteria to interpret the difference among them (Büyüköztürk, 2014; Kayış, 2014; Tezbaşaran, 2008).

In this context, item-total correlations were analyzed to reveal the distinctiveness capability of the items in the scale. Also, the significance of the difference between the average scores of the 27% sub (n= 242) and super (n= 242) groups was evaluated, which were established according to the scores related to the dimensions and total score. At this point, it was found as a result of the Pearson product-moment correlation coefficient analysis conducted for the item-total correlation that each item in the scale was significantly relevant to the total score of the scale positively (+) at .001 level (r> .30 and p< .001 for the 44 items). The correlation coefficient (r) defines a low-level relation if it is between .00 and .30 as an absolute value, a medium-level relation if it is between .30 and .70, a high-level relation if it is between .70 and 1.00. These results revealed that the items and dimensions in the scale illustrated similar behaviors, and the internal consistency of the trial was high (Büyüköztürk, 2014). According to these findings, it was observed that the item-total score correlation of the items in the scale was between .37 and .78, and item 20 (participates in group work actively) had the highest item correlation, while item 42 (is shy and introverted) had the lowest item correlation.

Besides, when the correlation between each dimension and the total score was examined, the scores were as follows: r= .89 in the "communication" dimension, r= .82 in the "adaptation" dimension, r= .70 in the "self-control" dimension, r= .86 in the "prosocial behaviors" dimension, and r= .86 in the "assertiveness" dimension (p <.001). These results revealed that each dimension in the scale was positively correlated with the total score (r> .30 and p> .001 for the five dimensions. In this scope, it was observed that the "communication," "adaptation,"

"self-control," and "prosocial behaviors" dimensions in the scale were highly correlated with (r> .70), while the "assertiveness" was correlated with the total score at a medium-level (r> .30). When the correlation of the dimensions with each other was examined, it was observed that all of the dimensions were significantly correlated with each other (p< .001), and the coefficients of this correlation varied between .33 and .89 (r> .30). This is also supported by the *t*-test independent groups results about the difference between the item mean scores of the 27% sub (n=242) and super (n=242) groups.

To reveal the distinctiveness capacity of the items and dimensions in the scale, these groups were formed according to the scores related to the dimensions and total items. As a result, it was determined that the mean scores of the super-group were significantly higher than the mean scores of the sub-group at the level of .001 (p <.001 for 44 items and five dimensions). At this point, Table 1 shows the t-test independent group results for each item. This result showed that the internal consistency of the trial was quite good, the internal consistency of the scale was ensured, and the items were distinctive for the children in terms of social skills. In addition, the Cronbach's alpha (α) reliability coefficient was also calculated to determine the reliability of the scale. It is recommended to use the Cronbach's alpha coefficient when there are more than three scores for items in a scale (Alpar, 2006). In this sense, as the scores obtained in the scale varied from one to five (1= never, 2= rarely, 3= sometimes, 4= often, and 5= always), the Cronbach's alpha reliability coefficient was considered to determine the reliability related to the dimensions and total scale.

When the Cronbach's alpha reliability coefficients of the scale were examined, the calculation results were as follows: α=.89 for the "communication", α= .90 for the "adaptation," α=.86 for the "self-control," α= .87 for the "prosocial behaviors", and α= .76 for the "assertiveness." The score for the total scale was α= .94. If .00 ≤ α < .40, it is not reliable; if .40 ≤ α < .60, it is low reliable; if .60 ≤ α < .80, it is quite reliable; if .80 ≤ α < .1.00, it is highly reliable for the criteria of the reliability based on the alpha (α) coefficient (Kayış, 2014). According to these values, it is clear that the scale is highly reliable for the "communication," "adaptation," "self-control," and "prosocial behaviors" dimensions and for the total scale, while it is quite reliable for the "assertiveness" dimension. The test-retest reliability was also considered another criterion for the reliability of the scale. The test-retest reliability reveals the consistency between the scores obtained from the scale at different times, and is determined according to the Pearson correlation results about the relationship between both sets of scores (Büyüköztürk, 2014). For the test-retest reliability of the scale, the teachers who previously had filled out the form were revisited four weeks later and asked to fill out the scale one more time for the same children for whom they had filled out the form before. It is recommended for this process that the time after the first practice should not exceed one month, as this period should not be too short (Aiken, 1997). In this context, the test-retest reliability was calculated using forms related to the 263 children, which were completely filled out and returned by the teachers.

The Pearson product-moment correlation coefficient was used to examine the relationship between the two sets of scores. The findings of the analysis revealed that the relationship between the two sets of scores was r= .82 in the "communication", r= .86 in the "adaptation," r= .75 in the self-control," r= .86 in the "prosocial behaviors", r= .63 in the "assertiveness", and r= .89 in the total (p <.001). In general, it would be safe to say that there was a positive and highly significant correlation in the dimensions and total scale. Nevertheless, the findings revealed that all of the items in the scale showed significant correlations between the scores of the two implementations (r> .70 and p <.001). This suggests that the test-retest reliability of the scale was sufficient and shows a stable construct depending on the time. At this point, the Cronbach's alpha reliability coefficient of the scale and the test-retest reliability analysis results indicated that the scale was reliable.

The findings of the study revealed that the psychometric values of the "SOSAS–TF" scale showed that the scale was a valid and reliable measurement instrument. Henceforth, a measurement instrument based on the assessment of the social skills of preschool children evaluated by the teachers has been developed; hence, the study made a contribution to the literature in this respect. Table 2 shows the descriptive information of the scale.

Table 2. Descriptive information on the SOSAS–TF scale

Dimensions	Scale item numbers	The number of items	The lowest value	The highest value
Communication	1-2-3-4-5-6-7-8-9-10-11-12	12	12.00	60.00
Adaptation	13-14-15-16-17-18-19-20-21	9	9.00	45.00
Self-control*	22*-23*-24*-25*-26*-27*-28*-29*-30*	9	9.00	45.00
Prosocial behaviors	31-32-33-34-35-36-37-38-39	9	9.00	45.00
Assertiveness*	40*-41*-42*-43*-44*	5	5.00	25.00
TOTAL		44	44.00	220.00

***Reverse scoring will be used for these items.*

Table 2 shows the descriptive characteristics of the SOSAS-TF scale. This scale consists of five dimensions and 44 items. While 30 of the items (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 31, 32, 33, 34, 35, 36, 37, 38, and 39) contain positive expressions, 14of them contain (22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 41, 42, 43, and 44) negative expressions. The dimensions of the SOSAS-TF scale have a minimum of five and a maximum of 12 items. The "communication" dimension has the maximum number of items (12 items), while the "assertiveness" dimension has the minimum number of items (5 items).

The "communication," "self-control," and "prosocial behaviors" dimensions in the scale have the same number of items (9 items). The social skills items in the scale were evaluated through a 5-level rating system according to how frequently the children exhibit them. Accordingly, when the status of a positively expressed social skill item to be exhibited by the children is "never," it was considered 1 point, "rarelyc 2 points, "sometimes" 3 points, "often" 4 points, and "always" 5 points. Reverse scoring (recoding) was applied for the negatively expressed items, when they were exhibited as "never" it was considered 5 points, "rarely" 4 points, "sometimes" 3 points, "often" 2 points, and "always" 1 point.

The probable score to be obtained from the dimensions in the scale is between 5 and 60. In addition, the probable score to be obtained from the total scale varies from 44 to 220. Evaluation of the scale takes an average of 15 minutes. The higher the scores obtained from the dimensions and the total, the higher the social skills. However, the lower the scores, the lower the social skills. It shows that there is a positive relationship between social skills and the scores of the scale.

Conclusions and Recommendations

In this study, the scale titled "Social Skills Assessment Scale-Teacher Form" for preschool children was developed and referred to as "SOSAS–TF". The findings of the study revealed valid and reliable the scale that was developed in the study is a valid and reliable measurement instrument. During the development of the scale (SOSAS–TF), teachers were consulted as the source person to evaluate the children's social skills. This scale consists of 44 items and five dimensions. These dimensions are named "communication," "adaptation," "self-control," "prosocial behaviors," and "assertiveness." The five factors of the scale explain 54.57% of the total variance. The factor loadings of all the items in the scale varied.46 and .74, and their common factor variances varied between .42 and .69. The item-total score correlations of the items in the scale varied between .37 and .78. Each dimension in the scale is positively and significantly correlated with the total score and with each other, and the coefficients of the correlation of the dimensions varied between .33 and .89.

According to the item average score comparison between 27% sub-super groups that was performed for the capability of distinctiveness of the items and dimensions in the scale, the score averages of the super-group were found to be significantly higher than the score averages of the sub-group. The Cronbach's alpha reliability coefficient (α) for the total scale was .94. The reliability coefficient for the dimensions in the scale is .89 for the "communication" dimension, .90 for the "adaptation" dimension, .86 for the "self-control" dimension, .87 for the "prosocial behaviors," and .76 for the "assertiveness" dimension.

The coefficient of test-retest reliability (r) for the total scale was .89. According to the test-retest reliability analysis performed for the dimensions in the scale the scores were as follows: r= .82 in the "communication" dimension, r= .86 in the "adaptation" dimension, r= .75 in the "self-control" dimension, r= .86 in the "prosocial behaviors" dimension, and r= .63 in the "assertiveness" dimension. The findings of the test-retest reliability for the dimensions and the total scale revealed that there is an overall positive and highly significant correlation. Thus, the results obtained in relation to the validity and reliability of the scale show that the scale can be used to evaluate preschool children's social skills.

The actions taken during the development of this scale were performed based on the data obtained from the preschool children. When the scale is used for groups other than preschool children, new validity and reliability studies with the data obtained from the relevant groups should be performed. Further studies can be conducted to test the validity and reliability of the SOSAS–TF scale for various age groups. Data for the children from the preschools and kindergarten classes in districts with different socioeconomic levels, in a city center in the Black Sea Region in Turkey, were obtained to increase the validity and ensure variety.

The study was conducted on the largest possible sample. Considering the sample size of the study, further comparative studies can be conducted for the validity and reliability of the scale in different regions and countries.

As an important step toward eliminating social skills deficiencies of children, it is recommended to operate processes for evaluating children’s existing social skills through a valid and reliable data collection instrument. The measurement instrument developed in this study can be used as a data collection instrument in descriptive or experimental studies for preschool children’s social skills.

The number of scales to measure social skills in Turkey and in the world is currently limited. Also, the SOSAS–TF scale will lead to filling a gap in the field. In this context, further studies having larger samples can measure the relationship between children’s social and other skills such as self-regulation, anger control, social adaptation, and communication to reassert the criterion validity of the scale. Studies on development of the social skills assessment scales for preschool children are considered to be useful not only for researchers and educators but also for decision makers. At this point, the information and findings about the social skills assessment scale can also be used to determine the effectiveness of intervention programs applied in educational institutions and their effects on children’s social-emotional development areas, including social skills.

The outcomes of this study affirm that the SOSAS–TF scale is a valid and reliable measurement instrument. To further extend its applicability, it is recommended that the scale be tested across different cultural and educational contexts. Future studies could explore the application of the scale in a broader context by examining its use in diverse regions, educational settings, and cultural backgrounds. Additionally, conducting longitudinal studies to assess the scale's stability over time would facilitate the tracking of changes in children's social skills, offering a more comprehensive understanding of their developmental processes.

Acknowledgements or Notes

This article is derived from the corresponding author’s doctoral dissertation entitled “Examination of the Effect of Story Telling Based and Play Based Social Skills Training on the Social Skills of Children Attending to Kindergarten,” supervised by the second author. The authors would like to express their sincere gratitude to the preschool teachers who supported the scale development process, as well as to the parents who kindly provided consent for their children's participation in the study. The authors also wish to declare that they received no financial support for this research.

Author(s) Contributions

The first and the second author were involved in designing the study. The first author was responsible for data collection, organizing the database, conducting the statistical analysis, and writing the original draft of the manuscript. The second author, as the supervisor, contributed to manuscript editing. Both authors reviewed and approved the final version of the manuscript.

Ethical Approval

Since the data for this study were collected before 2020, no Ethical Approval Form was obtained.

References

Aiken, L. R. (1997). *Psychological testing and assessment* (6th ed.). Allyn and Bacon.

Alpar, R. (2006). *Applied statistics in sports science*. Nobel Publications.

Arnold, D. H., Kupersmidt, J. B., Voegler-Lee, M. E., & Marshall N. (2012). The association between preschool children's social functioning and their emergent academic skills. *Early Childhood Research Quarterly, 27*(3), 376-386. https://doi.org/10.1016/j.ecresq.2011.12.009

Avcıoğlu, H. (2007). Reliability and validity study of the social skill evaluation scale (age 4-6). *Bolu Abant Izzet Baysal University Journal of Faculty of Education, 7*(2), 93-103.

Bacanlı, H. (2008). *Social skills training* (3rd ed.). Pegem Academy Publications.

Bandura, A. (1972). Modeling theory: Some traditions, trends, and disputes. In R. D. Parke (Ed.), *Recent trends in social learning theory* (pp. 35-62). Academic Press.

Bellini, S. (2008). *Building social relationships: A systematic approach to teaching social interaction skills to children and adolescents with autism spectrum disorders and other social difficulties.* Autism Asperger Publishing Company.

Boisjoli, J. A., & Matson, J. L. (2009). General methods of assessment. In Matson, J. L. (Ed.), *Social behavior and skills in children* (pp. 60-75). Springer Science+Business Media.

Büyüköztürk, Ş. (2014). *Data analysis handbook for social sciences: Statistics, research design, SPSS applications and interpretation* (19th ed.). Pegem A Publishing.

Campbell, C., Hansen, D. J., & Nangle, D. W. (2010). Social skills and psychological adjustment. In D. W. Nangle, D. J. Hansen, C. A. Erdley, & P. J. Norton (Eds.), *Practitioner's guide to empirically based measures of social skills* (pp. 51–67). Springer Publishing Company. https://doi.org/10.1007/978-1-4419-0609-0_4

Cerny, B. A., & Kaiser, H. F. (1977). A study of a measure of sampling adequacy for factor- analytic correlation matrices. *Multivariate Behavioral Research, 12*(1), 43-47. https://doi.org/10.1207/s15327906mbr1201_3

Ceylan, Ş. (2009). *Validity-reliability study of vineland social-emotional early childhood scale and examination of effects of creative drama education on social-emotional behavior of 5-year old pre-school children attending preschool educational institutions* [Unpublished doctoral thesis]. Gazi University.

Clayton, R. J., Hein, S., Keller-Margulis, M. A., & Gonzalez, J. E. (2021). Associations between social skills, inattention, and english vocabulary skills of preschool latinx dual language learners. *Journal of Research in Childhood Education*, *36*(2), 219–238. https://doi.org/10.1080/02568543.2021.1934200

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2th ed.). Lawrence Erlbaum Associates.

Conger, J. C., & Keane, A. P. (1981). Social skills intervention in the treatment of isolated or withdrawn children. *Psychological Bulletin, 90*(3), 478-495.

Connolly, J., & Doyle, A. B. (1981). Assessment of social competence in preschoolers: Teachers versus peers. *Developmental Psychology, 17*(4), 454–462. https://doi.org/10.1037/0012-1649.17.4.454

Cooke, T. & Apolloni, T. (1976). Developing positive social-emotional behaviors: A study of training and generalization effects. *Journal of Applied Behavior Analysis, 9*(1), 65-78. https://doi.org/10.1901/jaba.1976.9-65

Cornish, U. & Ross, F. (2004). *Social skills training for adolescents with general moderate learning difficulties.* Jessica Kingsley Publishers.

Çağlar, D. (1981). *Incompatible children and their education* (Extended 2nd ed.). Ankara University Printing House.

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2014). *Multivariate statistics SPSS and LISREL applications for social sciences* (3rd ed.). Pegem Academy Publications.

Deutsch, F. (1974). Observational and sociometric measures of peer popularity and their relationship to egocentric communication in female preschoolers. *Developmental Psychology, 10*(5), 745–747. https://doi.org/10.1037/h0037019

Dong, X., Burke, M. D., Ramirez, G., Xu, Z., & Bowman-Perrott, L. (2023). A meta-analysis of social skills interventions for preschoolers with or at risk of early emotional and behavioral problems. *Behavioral Sciences, 13*(11), 940. https://doi.org/10.3390/bs13110940

Elliott, S. N., & Busse, R. T. (1991). Social skills assessment and intervention with children and adolescents: Guidelines for assessment and training procedures. *School Psychology International, 12*(1-2), 63–83. https://doi.org/10.1177/0143034391121006

Erkul, R., & Sonmez, S. (2020). Relationship among social skills, problem behaviors and social competence of preschoolers. *i-manager's Journal on Educational Psychology*, 14(1), 49-61. https://doi.org/10.26634/jpsy.14.1.17189

Erkuş, E. (2012). *Measurement and scale development in psychology-I: Basic concepts and operations*. Pegem Academy Publications.

Erwin, P. G. (1994). Effectiveness of social skills training with children: A meta-analytic study. *Counselling Psychology Quarterly*, *7*(3), 305–310. https://doi.org/10.1080/09515079408254154

Finch, W. H. (2019). *Exploratory factor analysis.* SAGE Publications.

Foster, S. L., Inderbitzen, H. M., & Nangle, D. W. (1993). Assessing acceptance and social skills with peers in childhood: Current issues. *Behavior Modification, 17*(3), 255-286. https://doi.org/10.1177/01454455930173003

Gresham F. M. (2001). Assessment of social skills in children and adolescents. In J. Andrews, D. Saklofske, & H. Janzen (Eds.), *Handbook of psychoeducational assessment* (pp. 325–355). Academic Press.

Gresham, F. M. (1981). Social skills training with handicapped children: A review. *Review of Educational Research, 51*(1), 139-176. https://doi.org/10.2307/1170253

Gresham, F. M. (2002). Teaching social skills to high-risk children and youth: Preventive and remedial strategies. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches* (pp. 403–432). National Association of School Psychologists.

Gresham, F. M. (2016). Social skills assessment and intervention for children and youth. *Cambridge Journal of Education, 46*(3), 319-332. https://doi.org/10.1080/0305764X.2016.1195788

Gresham, F. M. & Elliott, S. N. (1984). Assessment and classification of children's social skills: A review of methods and issues. *School Psychology Review, 13,* 292-301.

Gresham, F. M. & Elliott, S. N. (1987). The relationship between adaptive behavior and social skills: Issues in definition and assessment. *The Journal of Special Education, 21*(1), 167-181. https://doi.org/10.1177/002246698702100115

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system.* American Guidance Service.

Gresham, F. M., Van, M. B., & Cook, C. R. (2006). Social skills training for teaching replacement behaviors: remediating acquisition deficits in at-risk students. *Behavioral Disorders, 31*(4), 363-377. https://doi.org/10.1177/019874290603100402

Gülay, H. (2004). *Comparing with the social skills of children living with their parents and orphans who are six years old* [Unpublished master's thesis]. Hacettepe University.

Gülay, H., & Akman, B. (2009). *Social skills in preschool period.* Pegem Academy Publications.

Günindi, Y. (2023). The effect of online education on children's social skills during the covid-19 pandemic. *International Electronic Journal of Elementary Education, 14*(5), 657–665. https://orcid.org/0000-0001-5192-5466

Hosp, J. L., Howell, K. W., Hosp, M. K. (2003). Characteristics of behavior rating scales: implications for practice in assessment and behavioral support. *Journal of Positive Behavior Interventions, 5*(4), 201-208. https://doi.org/10.1177/10983007030050040301

Jones, W. H., Hobbs, S. A., & Hockenbury, D. (1982). Loneliness and social skill deficits. *Journal of Personality and Social Psychology, 42*(4), 682–689. https://doi.org/10.1037/0022-3514.42.4.682

Kalaycı, Ş. (2014). Factor analysis. In Kalaycı, Ş. (Ed.), *SPSS applied multivariate statistical techniques* (pp. 321-345), (6th ed.). Asil Publication Distribution.

Kalutskaya, I., Archbell, K. A., Rudasill, K. M., & Coplan, R. J. (2015). Shy children in the classroom: From research to educational practice. *Translational Issues in Psychological Science, 1*(2), 149-157. https://doi.org/10.1037/tps0000024

Kamaraj, I. (2004). *The adaptation teacher form and parent form of preschool period (4-6 years old) social skill measurement scale into Turkish and the effect of educational drama program in acquiring assertive social skills of five years old children* [Unpublished doctoral thesis]. Marmara University.

Kapıkıran, N., İvrendi, A. B., & Adak, A. (2006). Social skills in preschool children: Detection of the situation. *Pamukkale University Journal of Education, 19*(1), 20-28.

Kavale, K. A., Forness, S. R., & Bender, M. (1988). *Handbook of learning disabilities: Methods and interventions*. Little Brown and Company.

Kayış, A. (2014). Reliability analysis. In Kalaycı, Ş. (Ed.), *SPSS applied multivariate statistical techniques* (pp. 403-419), (6th ed.). Asil Publication Distribution.

Kılıç, M., Tutkun, C., & Tezel-Şahin, F. (2023). The mediating role of social skills in the relationship between preschool children's academic self-esteem and problem behaviours. *Journal of Innovative Research in Teacher Education, 4*(3), 393-409. https://doi.org/10.29329/jirte.2023.616.3

Kline, P. (2014). *An easy guide to factor analysis.* Routledge.

Koçyiğit, S., & Kayılı, G. (2008). Comparison of the social skills of kindergarten students who received and did not receive Montessori education. *The Journal of Selcuk University Social Sciences Institute, 20*, 511-516.

Kutlu, M., & Kaya, A. (2005). Personal guidance and psychological counseling. In Kaya, A. (Ed.), *Psychological advice and guidance* (pp. 203-320) (2nd ed.). Anı Publications.

Leary, M. R., & Kowalski, R. M. (1995). *Social anxiety.* The Guilford Press.

Liberman, R. P., DeRisi, W. J., & Mueser, K. T. (1989). *Psychology practitioner guidebooks: Social skills training for psychiatric patients.* Pergamon Press.

Lodder, G. M. A., Goossens, L., Scholte, R. H. J., Engels, R. C. M. E., & Verhagen, M. (2016). Adolescent loneliness and social skills: Agreement and discrepancies between self-, meta-, and peer-evaluations. *Journal of Youth and Adolescence, 45*(12), 2406-2416. https://doi.org/10.1007/s10964-016-0461-y

Malecki, C. K., & Elliot, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17*(1), 1–23. https://doi.org/10.1521/scpq.17.1.1.19902

McDonald, R. P. (2014). *Factor analysis and related methods.* Lawrence Erlbaum Associates, Inc.

McFall, R. M. (1982). A review and reformulation of the concept of social skills. *Behavioral Assessment, 4*(1), 1-33. https://doi.org/10.1007/BF01321377

Merrell, K. W. (1996). Social-emotional assessment in early childhood. The preschool and kindergarten behavior scales. *Journal of Early Intervention, 20*(2), 132-145. https://doi.org/10.1177/105381519602000205

Merrell, K. W. (1999). *Behavioral, social, and emotional assessment of children and adolescents*. Lawrence Erlbaum Associates, Inc.

Merrell, K. W. (2001). Assessment of children's social skills: Recent developments, best practices, and new directions. *Exceptionality, 9*(1&2), 3–18. https://doi.org/10.1207/S15327035EX091&2_2

Merrell, K. W., & Gimpel, G. (2014). *Social skills of children and adolescents: Conceptualization, assessment, treatment.* Psychology Press.

Meuwissen, A. S. (2022). Social skill development of young children amid the pandemic. *Zero to Three*, 42(3), 37-43.

Ministry of National Education [MoNE] (2013). *Okul öncesi eğitim programı [Preschool education program]*. Ministry of National Education.

Mulvey, N., Jenkins, L., & Jenkins, L. (2020). Language skills as predictors of social skills and behaviors in preschool children. *Contemporary School Psychology*. https://doi.org/10.1007/S40688-020-00281-1

Offer-Boljahn, H., Hövel, D. C., & Hennemann, T. (2022). Learning behaviors, executive functions, and social skills: A meta-analysis on the factors influencing learning development in the transition from kindergarten to elementary school. *Journal of Pedagogical Research, 6*(1), 1-17. https://dx.doi.org/10.33902/JPR.20221175398

Parker, G., & Asher, S. R. (1987). Peer relations and later personal adjustment: Are low-accepted children at risk?. *Psychological Bulletin, 102*(3), 357-389. http://dx.doi.org/10.1037/0033-2909.102.3.357

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Sage Publication.

Quinn, M. M., Kavale, K. A., Mathur, S. R., Rutherford, R. B., & Forness, S. R. (1999). A meta-analysis of social skill interventions for students with emotional or behavioral disorders. *Journal of Emotional and Behavioral Disorders, 7*(1), 54-64. https://doi.org/10.1177/106342669900700106

Rademacher, A., Zumbach, J., & Koglin, U. (2020). Cross-lagged effects of self-regulation skills and behaviour problems in the transition from preschool to elementary school. *Early Child Development and Care, 192*(4), 631–637. https://doi.org/10.1080/03004430.2020.1784891

Rotter, J. B. (1982). *The development and applications of social learning theory.* Praeger.

Scherer, R. F., Wiebe F. A., Luther, D. C., & Adams, J. S. (1988). Dimensionality of coping: Factor stability using the ways of coping questionnaire. *Psychological Reports, 62*(3), 763-770. http://dx.doi.org/10.2466/pr0.1988.62.

Schneider, B. H. & Byrne, B.M. (1985). Children's social skills training: A meta-analysis. In B. H. Schneider, K. H. Rubin, & J. E. Ledingham (Eds), *Children's peer relations: Issues in assessment and internetin* (pp. 175-192). Springer. https://doi.org/10.1007/978-1-4684-6325-5_11

Segrin, C. (2000). Social skills deficits associated with depression. *Clinical Psychology Review, 20*(3), 379-403. http://dx.doi.org/10.1016/S0272.

Segrin, C., & Flora, J. (2006). Poor social skills are a vulnerability factor in the development of psychosocial problems. *Human Communication Research, 26*(3), 489-514. https://doi.org/10.1111/j.1468-2958.2000.tb00766.x.

Sharma, R., Goswami, V., & Gupta, P. (2016). Social skills: Their impact on academic achievement and other aspects of life. *International Journal for Innovative Research in Multidisciplinary Field, 2*(7), 219-224.

Sharma, S. (1995). *Applied multivariate techniques*. John Wiles & Sons, Inc.

Sheridan, S.M., & Walker, D. (1999). Social skills in context: Considerations for assessment, intervention, and generalization. In C. R. Reynolds & T. B. Gutkin (Eds.), *The handbook of school psychology* (3rd ed.), (pp. 686–708). Wiley & Sons.

Strain, P. S., Odom, S. L., & McConnell, S. (1984). Promoting social reciprocity of exceptional children: Identification, target behavior selection, and intervention. *Remedial and Special Education, 5*(1), 21-28.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson Publications.

Takahashi, Y., Okada, K., Hoshino, T., & Anme, T. (2015). Developmental trajectories of social skills during early childhood and links to parenting practices in a Japanese sample. *PLoS ONE, 10*(8), e0135357. https://doi.org/10.1371/journal.pone.0135357

Tavşancıl, E. (2005). *Measuring attitudes and data analysis with SPSS* (2nd ed.). Nobel Publishing Distribution.

Tezbaşaran, A. A. (2008). *Likert type scale preparation guide (e-book)* (3rd version). Turkish Psychological Association Publications.

Warnes, E. D., Sheridan, S. M., Geske, J., & Warnes, W. A. (2005). A contextual approach tothe assessment of social skills: Identifyingmeaningful behaviors for social competence. *Psychology in the Schools, 42*(2), 173–187. http://dx.doi.org/10.1002/pits.20052

Welsh, M., Parke, R. D., Widaman, K., & O'Neil, R. (2001). Linkages between children's social and academic competence: A longitudinal analysis. *Journal of School Psychology, 39*(6), 463–482. https://doi.org/10.1016/S0022-4405(01)00084-X

Whirter, J. M. & Voltan Acar, N. (1998). *Communication with the child.* National Education Press.

Wilson, M. J., & Bullock, L. M. (1989). Psychometric characteristics of behavior rating scales: Definitions, problems, and solutions. *Behavior Disorders, 14*(3), 186-200. https://doi.org/10.1177/019874298901400305

Zöggeler-Burkhardt, L., Embacher, E. M., & Smidt, W. (2023). Social relationships, interactions and learning in early childhood – theoretical approaches, empirical findings and challenges. *Early Child Development and Care, 193*(11–12), 1199–1203. https://doi.org/10.1080/03004430.2023.2260976

Zsolnai, A. (2002). Relationship between children’s social competence, learning motivation and school achievement. *Educational Psychology, 22*(3), 317–329. https://doi.org/10.1080/01443410220138548