Article

The Social and Emotional Learning and Orientation Scale - Development and Validation of a Brief Measure in Hindi

Journal of Psychoeducational Assessment
2022, Vol. 40(5) 571–591

DOI: 10.1177/07342829221075517
journals.sagepub.com/home/jpa
SAGE

Mayank Sharma[1], Sonal Chheda[1], Reshma Piramal[2], Neha Bhatia[2], Tyralynn Frazier[3], and Nandini Chatterjee Singh[1]

Abstract

Accumulating evidence of the impact of social and emotional learning (SEL) on student well-being as well as academic performance has led to an increased interest in understanding and measuring SEL in students. However, most existing measures have been developed in English and there is a paucity of culturally relevant measures in other languages. In this study, we present a short, culturally sensitive 8-item scale in Hindi entitled the Social and Emotional Learning and Orientation Scale (SELOS). Factor structure of SELOS was established through exploratory (N=2844; M_{age}=14.65 years) and confirmatory factor analysis (N=1508; M_{age}=14.50 years) with responses from 4352 Hindi-speaking students in India. The scale demonstrated a two-factor structure, with good reliability, discriminant validity and partial measurement invariance across gender. These results provide encouraging evidence for conducting culturally sensitive studies of SEL in India and in other Hindi-speaking parts of the world.

Keywords

social and emotional learning, Hindi, adolescent, survey, India

SEL in India

India has one of the highest adolescent and child populations in the world (approximately 444 million) (UNICEF, 2020). The 2016 National Mental Health Survey conducted by the Government of India found that 7.3% of children aged 13–17 years suffered from or were at the risk of

[1]UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development, New Delhi – 110 001, India
[2]Social, Emotional and Ethical Learning India, Max India Foundation, Noida, India
[3]Center for Contemplative Science and Compassion-Based Ethics (CCSCBE) Atlanta, GA, USA

Corresponding Author:
Nandini Chatterjee Singh Full address: UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development, First Floor, Copernicus Marg, 35, Ferozeshah Road, Connaught Place, New Delhi – 110 001, India
Email: n.chatterjee@unesco.org

a mental health disorder (Gururaj et al., 2016). Results from a recent WHO report also revealed that among 13–15 year old Indian adolescents, 25% showed signs of depression, 8% showed signs of loneliness and anxiety, 10% said they have "no close friends," and 7% reported experiencing some sort of bullying in the classroom (WHO, 2017). This indicates that a sizable population of children and adolescents face mental well-being issues and have problems in creating and maintaining peer relationships. Moreover, these statistics show prevalence of several mental health issues among learners in India and warrant interventions that put learner well-being at the forefront to create conducive learning environments.

Accumulating empirical evidence supporting the impact of Social and Emotional Learning (SEL) in improving the well-being and mental health of students has made SEL prominent in many classrooms today. For instance, at least two meta-analyses investigating the efficacy of SEL programs on child and adolescent development have revealed significant effect sizes for: a) decreased externalized behaviors such as lack of compliance, aggression, bullying, and delinquency (Boncu et al., 2017) and b) positive impact on self-regulation, problem solving, and relationship skills (Taylor et al., 2017). Both these analyses also indicate that all of these skills described above contribute to increased positive attitudes towards the self and towards others, as well as address mental health problems such as emotional distress and conduct disorders (Durlak et al., 2011). Additionally, in a review of 317 studies conducted by Payton et al. (2008) involving 324,303 elementary and middle school children showed that SEL interventions added 11 to 17 percentile point gains to student academic achievement tests. Taken together, these studies show that the introduction of SEL addresses aspects of relationship and emotion building in order to support a better and brighter future for students (Elias et al., 1997).

SEL may be broadly described as the process of recognizing and managing thoughts, feelings, and behaviors in order to become aware of the self and of others, make responsible decisions, and manage one's behaviors (Zins & Elias, 2007). SEL has gained traction due to research which states that socially connected, and emotionally stable children not only perform better academically but also go on to become productive, socially competent and responsible citizens needed for the 21st century (Denham et al., 2016; Oberle et al., 2016). Moreover, recent research from cognitive development showing that social and emotional skills can be taught in the classroom in a manner similar to other forms of academic learning (Masten, 2001) has attracted the attention of various stakeholders towards SEL.

SEL has been present in Indian curricula in disparate forms owing to the broad aims of the National Council for Educational Research and Training (2005). However, the new National Education Policy released in 2020 (Ministry of Human Resource Development, Government of India, 2020) has made SEL goals explicit and can be considered a move towards mainstreaming it in Indian schools. Although this is welcome news, it also presents challenges for effective implementation, particularly because there are no culturally sensitive scales available in Indian languages that can be used to assess needs of learners and design appropriate interventions.

Existing Measures of SEL

The current landscape of tools that are available to assess social and emotional competence is broad. Existing measures have originated from multiple frameworks catering to topics such as youth risk, developmental assets and mental health. One popular measure in the field of SEL is the Behavior Assessment System for Children - 3 (BASC-3; Reynolds & Kamphaus, 2015), which is a battery of instruments that help in comprehensive assessment of behavioral and emotional functioning of children, adolescents, and young adults. A second popular example is the School Social Behavior Scales - 2 (SSBS - 2; Merrell, 2002), which is a measure that is used to assess social competence and antisocial behavior in students.

In a detailed review conducted by Humphrey et al. (2011), more than 189 such measures were identified. Out of these, the majority (172 out of 189) were either: a) lengthy (took more than 30 min to complete), b) did not allow self-report by students, c) had open ended responses which required manual coding, d) measured only one particular construct of social and emotional competence instead of both or, e) had not been psychometrically validated.

The use of short scales in conducting psychological assessment has become very popular (Kruyen et al., 2013). The main advantage of using short scales is the saving of assessment time (Morgado et al., 2017). Additionally, using shorter assessments reduces costs and improves participation rates in studies (Edwards et al., 2004). They also help avoid fatigue, which might reduce the quality of collected data. However, the use of shorter scales is often critiqued in the literature because they may lead to a reduction in its psychometric properties (Credé et al., 2012). But at the same time, emerging data also shows that if one considers appropriate statistics, the development of shorter scales can aid the quality of data collected and also ensure respondent motivation. For instance, Kemper et al. (2018) suggest the use of more appropriate statistics such as polychoric correlations instead of Pearson correlations for responses with fewer Likert scale options, or use of item response analysis coupled with factor analysis for scales with a lesser number of items. In fact, they also suggest that after weighing in on factors such as use of appropriate statistics, rate and quality of data collection, and psychometric evidence supporting the intended use of the scale, an appropriate choice between short and long scales can be made (Kemper et al., 2018).

A few brief SEL scales have been developed, such as the SSIS SEL*b*-S (Anthony et al., 2021) and the short form version of the SEARS for adolescents (SEARS-A SF; Nese et al., 2012). The SSIS SEL*b*-S was validated on data obtained from students in Grades 3–12 and was meant to overcome challenges faced while using long-form measures such as large-scale (e.g., universal assessment/screening) or repeated (e.g., periodic progress monitoring) assessment. Similarly, the SEARS-A SF was developed to serve the purposes of systematic progress monitoring and screening. When using the statistical methods described above, both scales (the SSIS SEL*b*-S and the SEARS-A SF) showed good psychometric properties and the SEARS-A SF also showed excellent correlational validity with the long-form version (Anthony et al., 2021; Nese et al., 2012).

However, even when using brief scales, an important issue faced with existing measures is that of cultural relevance and challenges in translation. For instance, many scales use proxy indicators for evaluation of SEL like questions related to school attendance, which are unfortunately not culturally suited for Indian populations. For example, in a developing country like India, absence of attendance can be linked to various reasons such as ill-health, contribution to farm labor during harvest seasons, seasonal migration, or even difficulty in access. Thus, attendance cannot provide reliable information on social and emotional orientations or states of learners, particularly in semi-urban or rural areas of India. In addition to proxy indicators, language and translation issues pose a major challenge in administering lengthy and complex SEL scales among such populations. Translation into local languages is challenging not only because cultural transferability cannot be implicitly assumed by translation but also because of the varying structures and syntax of different languages themselves (Cha et al., 2007). For one, this presents challenges in finding appropriate substitute words in local languages for concepts that may be completely foreign. For instance, concepts such as "hanging out," which are part of items in many western scales, including the SEARS-A, do not have an equivalent term in the local Indian context. Similarly, a concept such as whiling away time with friends might have a negative connotation attached to them, especially in rural areas of India where children are often active contributors in household chores. Consequently, when translated, tools may not fully capture values inculcated through local cultural and social practices (Wang et al., 2006).

The absence of culturally relevant, linguistically appropriate tools is also reflected in research. A review of the SEL literature shows that, much of the published research in SEL has been conducted in North America (e.g., Rimm-Kaufmann et al., 2014; Schonert-Reichl et al., 2015), Europe, and the United Kingdom (e.g., Banerjee et al., 2013). In comparison, few studies have been conducted in Asia. In fact, Torrente et al.'s (2015) mapping of SEL policies in Asia revealed that SEL has gained prominence over the years to facilitate a shift from a heavy academic focus of Asian education systems to include more 21st century skills. It is worthwhile to acknowledge that this assumes an implicit prevalence of SEL competencies in national goals of education with an underlying focus on cultural consonance in these societies (Lagi & Armstrong, 2017), but explicit measurement of SEL has still been absent. It is likely that this lack of measurement is due to the paucity of culturally appropriate measurement tools. Because of the large-scale implementation of potential evaluations in India and the concurrent translation challenges, it is necessary to develop brief, easy to administer, and culturally sensitive self-report questionnaires in local languages that adequately provide educators a snapshot of children's social and emotional states.

Purpose

The aim of this study was to design a brief, culturally sensitive, psychometrically validated questionnaire in Hindi to ascertain SEL states of students in India. As per the Census of India conducted in 2011, Hindi is not only one of the two official languages of India but also the most widely spoken language with close to half a billion speakers (Census of India, 2018; Government of India, 2018). As of 2019, close to 42% of all learners in the country were enrolled in schools where Hindi is the primary language of instruction (Unified District Information System for Education, 2019). It is therefore necessary that any tool developed for SEL measurement in India be first developed in Hindi, followed by other local languages of instruction.

Based on the research described above, it was decided that a brief scale would be developed in Hindi to obtain some baseline measurement of social and emotional states of children and adolescents. It was also postulated that results from such a baseline measurement would help highlight to policymakers the urgent need for SEL in schools and curricula in India. Given the advent of COVID-19 that has led to school closures, the study was conducted online where students took the assessment at home. The assessment was unsupervised (without the presence of teachers) and taken on mobile devices. Because of concerns around internet fatigue due to online teaching and collection of spurious data from lengthy questionnaires that might put off students, a brief scale was most appropriate for this study.

Method

Participants

Data was collected in two phases. In the first phase of data collection, a total of 2844 students completed the questionnaire. Participants from school grades 3–12 took part in the study and their grade distribution was as follows: Grades 3–5 (4.5%), Grades 6–8 (20.3%) and Grades 9–12 (75.2%). They reported a mean age of 14.65 years (SD = 2.11 years, range 9–18 years).

In the second round of data collection, 1508 students completed the questionnaire. Grade distribution of participants was as follows: Grades 3–5 (5.2%), Grades 6–8 (22%) and Grades 9–12 (72.8%). Participants reported a mean age of 14.50 years (SD = 2.08 years, range 9–18 years). Out of these respondents, 60.54% students reported their gender as female, while 39.46% reported their gender as male. Females (N = 913) reported a mean age of 14.41 years (SD = 1.98 years, range 9–18 years) and males (N = 595) reported a mean age of 14.65 years (SD = 2.20 years, range

9–18 years). An unpaired *t*-test revealed a significant difference ($p < .05$) in the ages of male and female respondents, with males reporting a significantly higher average age than females.

Item Development

The process of item development followed five steps. In the first step, criteria for item development were laid out. Drawing from literature, three important criteria were established in order to create a scale that could address existing gaps in SEL literature, suit the cultural needs of a Hindi-speaking Indian student population, and ensure quality collection of data. These included brevity and simplicity of scale items, inclusion of items that were culturally relevant and that measured both social and emotional competence (based on the shortcomings of existing SEL measures pointed out by Humphrey et al., 2011). The WHO and CDC designed Global School-based Student Health Survey (GSHS; WHO, *n*. d.) was used as a reference to create the items on social and emotional competence. In the second step, a review of existing measures of SEL was performed. Popular SEL scales were reviewed and their items were examined from the viewpoint of cultural sensitivity and brevity. In the third step, a pool of items was created by one of the authors of the study who was a native speaker of Hindi and had over 20 years of research experience in the fields of cognitive neuroscience, psychology, SEL, as well as scale development in Indian languages. A total of 11 items were created as it was envisioned that after item reduction through factor analysis, the final scale would consist of 6–8 items. Important item development guidelines such as avoiding complex language, ambiguous questions, double-barreled items, jargon and biased questions were followed to reduce measurement error (Devellis, 2012). Cultural sensitivity of items was ensured by avoiding cognitive and verbal lexicon of social and emotional experiences used in Western measures (such as "feeling down", "getting swept up in feelings", etc.) since they have no literal translations in Hindi, and cannot be interpreted by Indian audiences. Based on the results of the implementation of the first cycle of the GSHS in India in the school years of 2006-07, items related to loneliness, relationships and bullying were included under social competence. As equivalents of these labels in Hindi can appear stigmatizing, especially for cultures like India where mental health remains a sensitive topic, they were avoided and softer alternatives were used to obtain accurate responses (for example, "*I prefer to be alone rather than with people my age*" instead of "*I experience loneliness*"). For gauging emotional competence, items related to experiencing basic emotions such as sadness, unhappiness, anger, etc. were included. Since Hindi follows a system of gendered verbs, verb options for both genders participating in the study (females and males) were indicated in the items using a slash ("/"). Finally, given the possible use of this scale with varying populations in India - urban, semi-urban and rural - items did not make any implicit or explicit assumptions about the socio-economic status, class, and access to basic amenities (except education) of the participants. Additionally, almost half the items (5 out of 11) were worded negatively in order to eliminate agreement bias that might appear while responding to the scale.

In the fourth step, expert review was conducted through consultation with a separate team of five members consisting of data analysts, SEL researchers and past teachers who had direct experience in working with students and designing age-appropriate digital SEL interventions. They were all native speakers of Hindi too. All of them provided feedback on the conciseness, clarity and completeness of items. Teachers working at the State Council of Educational Research and Training (SCERT) in Rajasthan also gave their feedback to ensure that the scale items were well-suited for administration in rural regions (SCERTs are academic bodies created by the State Departments for Education with a mission and objective to raise the quality of school education by improving the attitudes, application of knowledge and teaching skills of teachers). After

incorporating feedback from all sources, the items were examined again to reduce redundancy and increase ease of understanding.

In the fifth step, content validity of the scale was examined. Two members of the team independently mapped all 11 items of the scale (see complete list of items in Table 1) into the two categories of: a) social, and b) emotional competence. The mappings were then placed side by side and discussed. There was a near perfect agreement on this classification (1 item out of 11 indicated discrepancy which was resolved after discussion). The exercise indicated that five items on the scale measured emotional competence, while six measured social competence, which served as a hypothesized factor structure for the scale. However, the possibility of another factor structure based on separation of items according to polarity was not ruled out.

Response Format

All 11 items were to be responded to on a three-point Likert-type scale. Chambers and Johnston (2002) showed that younger children had greater difficulty in answering questions on abstract concepts such as their internal states than when answering questions based on concrete physical tasks. In addition, Wright and Asmundson (2003) found that simplifying existing five-point Likert-type scale questionnaires into three-point Likert-type scales made them easily

Table 1. Factor loadings for exploratory factor analysis (N = 2844).

Item	F1: Social and Emotional Orientations	F2: Social and Emotional Challenges
SELOS 1 - मुझे लोगों के साथ दोसती करना आसान लगता है। (I find it easy to make friends and be with people)	0.29	-0.08
SELOS 2 - मैं अकसर उदास और नाखुश रहता/रहती हूँ। (I experience sadness and unhappiness often)	0.01	**0.67**
SELOS 3 - मुझे अकसर गुससा आता है। (I often get angry and lose my temper)	-0.02	**0.51**
SELOS 4 - मैं अपने हमउम् साथियों के साथ रहने की बजाये अकेले रहना पसंद करता/करती हूँ। (I prefer to be alone rather than with people my age)	-0.03	**0.41**
SELOS 5 - मुझे एकागता से काय करने में कठिनाई होती है। (I find it difficult to concentrate on my studies/work at hand)	-0.09	0.36
SELOS 6 - मैं कुछ करने से पहले सोचता/सोचती हूँ। (I think before I do things)	**0.52**	-0.08
SELOS 7 - मेरे परिवार के लोग और सहपाठी मेरी भावनाओं को आहत करते हैं। (others around me bully me)	0.17	0.23
SELOS 8 - मैं सदैव दूसरों की भावनाओं का धयान रखता/रखती हूँ। (I pay attention to the feelings of others)	**0.74**	0.06
SELOS 9 - मैं अपने शि्धकों से सहजता से बात कर सकता/सकती हूँ। (I talk easily with teachers)	**0.49**	-0.02
SELOS 10 - मुझे दूसरों की मदद करना अ्छा लगता है। (I like to help others)	**0.73**	-0.05
SELOS 11 - मुझे लगता है कि सीखने में भावनाएँ मायने रखती हैं। (I think feelings matter in learning)	**0.45**	0.05

Note. Loadings forming each factor highlighted in bold (cutoff for factor extraction = 0.40). English translations for the Hindi items have been provided for the purpose of understanding the items, it should be noted that only the Hindi version of SELOS has been validated in this study.

understandable by children up to 15 years of age. Existing research has also suggested that most children have difficulties responding to Likert-type scale questions when the response formats are based on numbers (Moore & Mellor, 2003). A better way appears to be using words that indicate frequency of thoughts (for example, "Never" to "Always"). Additionally, Mellor and Moore (2013) found that older children tended to give more indefinite mid-scale responses (such as "Uncertain" or "Neutral") than their younger counterparts, suggesting that having more points on the response scales may be counterproductive in gauging the true internal states of children. At the same time, a dichotomous scale may be too narrow and definitive in capturing their responses. Since our questionnaire was developed to be administered across the 9–18 age group, it was more prudent to limit the response to a three-point Likert-type scale to ensure it was inclusive of the varied developmental stages of each respondent. The response anchors for the three-point Likert-type scale provided to the participants read "Never", "Sometimes" and "Always" (in Hindi) and avoided using indefinite mid-points as well as number-based formats.

Procedure

The questionnaire was approved and delivered virtually to government school students in Rajasthan, India by the State Council for Educational Research and Training (SCERT), via SEE Learning India's partnership with Boston Consulting Group (BCG). Rajasthan is among the 10 most populous states in India and has the second highest number of Hindi speakers in the country, with 89% of its populace reporting they either speak Hindi or one of its dialects (Government of India, 2018). There is the existence of local regional languages spoken in Rajasthan besides Hindi. However, all of them belong to the same Indo-European group of languages that are mostly intelligible to Hindi speakers. During the country-wide lockdown due to the COVID-19 pandemic, students lost access to schools overnight and teachers and educators received concerns about the mental well-being of children. Thus, BCG partnered with SEE Learning India to provide SEL content and evidence-based practices to all children through the SCERT. Along with this intervention, a short survey (SELOS) was developed to ascertain insights into the social and emotional competence of the children.

Since the intervention and the survey were part of a partnership between SEE Learning India, BCG and the SCERT, all students, ranging from grades 3 through 12, from the roughly 11,000 senior secondary schools in the state (Department of Education, State Government of Rajasthan, 2018) were eligible for the study. Permission to conduct the study and collect the data were provided by the SCERT (the state government body who helped identify schools, administer the survey, and conduct the study with protocol). Participation in the study was voluntary and completely anonymous. No strict measures of completion were put into place since the responses could not be tracked to individual students due to anonymity. Collection of sensitive information such as ethnicity and religion was strictly avoided, however, information on the district of the participants was collected in order to ensure representation, as well as ensure that the predominantly Hindi-speaking populace was included in the study.

Participating students in the study received formal school education in Hindi. As the study was conducted during the COVID-19 school closures, students participated remotely from their homes. Using a cascading model of WhatsApp group message chains, the SCERT first shared the scale with the schoolteachers, who forwarded it to their students' parents and guardians through WhatsApp groups. The students were instructed to take the questionnaire individually, and at their ease. They were briefed about the purpose of the scale and could get their queries resolved through their teachers via message or call. The scale was administered virtually through Google Forms and required 3–4 min to complete. Students were given 7 days to complete the form.

The 11-item scale was delivered to students during the first round of data collection. In addition to completing this scale, participants also completed a demographic questionnaire to collect data on their age, grade, and district. Following this, Exploratory Factor Analysis (EFA) was conducted on this data. This led to a revised 8-item scale, that was delivered virtually to a new sample of government school students in Rajasthan. These data were then used to perform a confirmatory factor analysis (CFA). In this phase, additional information on the gender of participants was collected using the existing demographic questionnaire.

Data Analysis

All data analyses were performed on R version 4.0.2 (https://www.r-project.org). On the data collected in the first phase ($N = 2844$), bivariate associations of scale items were first examined. Exploratory factor analysis was performed, and factors were extracted using results from commonly-used criteria: K1 rule (or the eigenvalue greater than one rule; Kaiser, 1960), Cattell's scree test (retaining eigenvalues below the point of inflection on the scree plot; Cattell, 1966), Horn's parallel analysis (comparing eigenvalues generated from the data matrix to the eigenvalues generated from a Monte-Carlo simulated matrix of random data of the same size; Horn, 1965), and Velicer's Minimum Average Partial (MAP) test (Velicer, 1976; Velicer et al., 2000). The hypothesised factor structure obtained from EFA was fitted on the data collected during the second phase ($N = 1508$) to perform a confirmatory factor analysis. Goodness of fit for the model was calculated using the following indices: the normed chi-square value ($\chi^2_{S\text{-}B}/df$) < 5 (Schumacker & Lomax, 2004), the robust CFI > 0.90 (Bentler, 1990), the robust GFI > 0.95 (Schermelleh-Engel et al., 2003), the robust RMSEA < 0.05 (Browne & Cudeck, 1992) and the robust SRMR < 0.08 (Hu & Bentler, 1999). Other psychometric properties of the scale such as reliability, discriminant validity and measurement invariance across genders were examined. Finally, factor scores were compared for differences across age and gender. All factor analytic procedures used polychoric correlations.

Results

Bivariate associations of scale items

Mean scores of individual items ranged from 1.55 to 2.87. The polychoric correlation matrix of item responses revealed a few moderate but significant correlations (see Figure 1). SELOS items 2, 3, 4 and 5 showed negative correlations with items 1, 6, 8, 9, 10 and 11. SELOS-7 ("*Others around me bully me"*) showed weak positive correlations with SELOS items 1, 6, 8, 9, 10 and 11, instead of the expected negative correlations (based on the polarity-based hypothesized factor structure). A closer look at the item indicated that it had been worded poorly (the current wording gauged more of others' social and emotional capabilities rather than the respondent's). Perhaps a better wording would have been *"I get bullied by others."* and hence, the expected negative correlation was not seen.

Exploratory factor analysis

Prior to conducting an EFA, it's prerequisites were tested. A Kaiser-Meyer-Olkin (KMO) value of 0.76 indicated adequacy of sample size for factor analysis (KMO value must be >.60 according to the rules of Kaiser, 1974). Additionally, Bartlett's Test of Sphericity (Bartlett, 1954) showed a significant result ($\chi^2(df = 55) = 152.25$, $p < .001$). This indicated that the correlation matrix was factorisable.

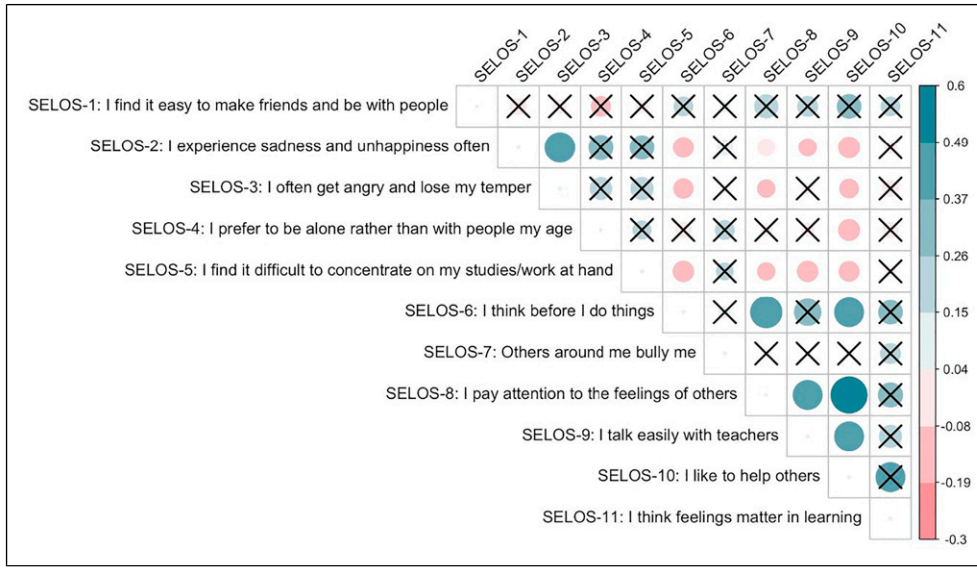

Figure 1. Polychoric correlation matrix for SELOS items (N = 2844) - crossed out correlations are not significant. Blue circles indicate positive associations and pink circles indicate negative associations. The intensity of colors and size of circles are indicative of the strength of correlation (auto-correlations have been excluded).

After prerequisite testing, questionnaire items were examined for univariate and multivariate normality. Results of the Shapiro-Wilk test (Shapiro & Wilk, 1965) for individual items indicated that they were not univariate normal (all p-values < .001*)*. In addition, results of the Henze-Zirkler test (Henze & Zirkler, 1990) (p < .001) indicated that the items were not multivariate normal either. Consequently, the Principal Axis Factoring (PAF) method was used for factor extraction as it is free from distributional assumptions (Fabrigar et al., 1999). In order to overcome the problem of factor indeterminacy (Velicer & Jackson, 1990), a Principal Component Analysis (PCA) was also conducted during the initial step for comparison. The solutions yielded through both the methods were similar, and thus the problem of factor indeterminacy was ruled out.

An unrestricted factor analysis of all 11 items in both PAF and PCA revealed three eigenvalues with magnitudes greater than one. Selection of the number of factors was done through comparison of results from the K1 rule, Cattell's scree plot, Horn's parallel analysis, and MAP test. The K1 rule indicated a 3-factor solution. The eigenvalues of these factors were 2.84, 1.52 and 1.07 respectively. Scree plot suggested a 2-factor solution, parallel analysis hinted at a 4-factor solution, and MAP test indicated a one factor solution. Keeping in mind the over and under extraction tendencies of these tests, it was hypothesised that the number of factors retained should be 2 or 3.

Gorsuch (1983) reported that in case of a confusion regarding retaining the number of factors, researchers should be biased towards selecting a greater number of factors. Therefore, a factor solution was obtained by extracting three factors from the data using a loading cutoff value of .40 (Tabachnick & Fidell, 2001). Direct oblimin rotation method was used to optimise the solution. The results of this solution showed that there was only one item (item 7) loading onto factor 2. It has been argued that each factor should have at least three items for it to be reliable (MacCallum et al., 1999). There were also no theoretical considerations that would place item 7 ("*Others around me bully me*") into a separate factor of its own (the item was either envisioned under the factor of "social competence" or the factor formed by items having negative polarity). Therefore, the 3-factor solution was

rejected, and a second factor solution was obtained by extracting two factors. This yielded a meaningful factor structure: SELOS - 6, 8, 9, 10 and 11 loaded onto factor 1 while SELOS - 2, 3, and 4 loaded onto factor 2 (details in Table 1). This was partially the same as the factor structure hypothesised based on polarity of items during item development. Multicollinearity was examined and ruled out since there were no cross-loading items present. Three items (SELOS - 1 and 7) did not load well onto any of the two factors. Some supplementary analysis was performed on these items using Item Response Theory (IRT) approaches. IRT can be applied to investigate which items do not have enough reliable information about the construct being measured and help differentiate item properties (such as discrimination and difficulty) among individuals. If the analyses show that there is such a problem with some items, the researcher can remove/modify those items and increase the validity of the scale (Oishi, 2007).

To conduct an item response analysis, Item Characteristic Curves (ICC) and Item Information Curves (IIC) were plotted using a Graded Response Model (GRM) for both polarity-based hypothesised factors (IRT approaches assume unidimensionality) to better understand how the participants responded to SELOS - 1, 5 and 7. Figures 2 and 3 represent the ICCs and IICs for factor 1 and 2, respectively. For both the graphs, theta (θ) represents a person's true latent trait (factor), that has been standardized to follow a normal distribution with a range from -3 to 3, where 0 represents the average score (Baker, 2001). For the ICC, $P(\theta)$ represents the probability of responding to specific categories (P1 = Never, P2 = Sometimes, and P3 = Always) on an item's response scale, while for the IIC, $I(\theta)$ represents the information function, or how well each item contributes to score estimation precision (higher levels of information leading to more accurate score estimates). In an ICC, discrimination is defined as how well an item can differentiate between examinees having abilities below the item location and those having abilities above the item location. This property reflects the steepness of the item characteristic curve in its middle section. The steeper the curve, the better the item can discriminate. Similarly, in an IIC, the amount of information an item contributes depends on its slope parameter - the larger the parameter, the

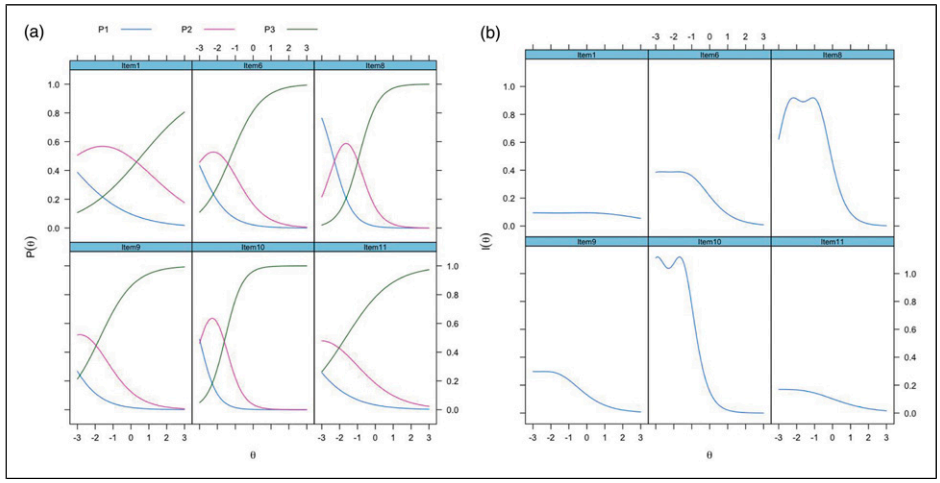

Figure 2. (a) Item Characteristic Curves (ICC), and (b) Item Information Curves (IIC) for SELOS items forming hypothesised factor 1 (SELOS - 1, 6, 8, 9, 10 and 11). For the ICC, $P(\theta)$ represents the probability of responding to specific categories (P1 = Never, P2 = Sometimes, and P3 = Always) on an item's response scale, while for the IIC, $I(\theta)$ represents the information function, or how well each item contributes to score estimation precision. θ refers to the respondent's ability to respond to an item, standardised in the range of -3 to 3.

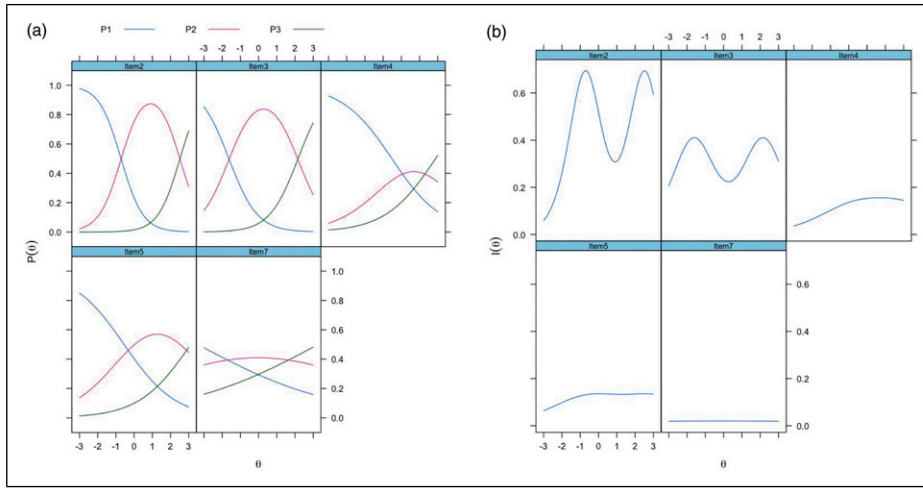

Figure 3. (a) Item Characteristic Curves (ICC), and (b) Item Information Curves (IIC) for SELOS items forming hypothesised factor 2 (SELOS - 2, 3, 4, 5, and 7). For the ICC, $P(\theta)$ represents the probability of responding to specific categories (P1 = Never, P2 = Sometimes, and P3 = Always) on an item's response scale, while for the IIC, $I(\theta)$ represents the information function, or how well each item contributes to score estimation precision. θ refers to the respondent's ability to respond to an item, standardised in the range of -3 to 3.

more information the item provides. As seen in Figure 2(a), item one had the worst level of discrimination as it had the least steep curve. Item one also had the lowest slope, and was therefore, the least informative item, as also demonstrated by the results of EFA (see Figure 2(b)). Similar trends in steepness of curves and values of slopes were observed for items 5 and 7 in Figures 3(a) and 3(b). Hence, the criteria for elimination of these items were strengthened since they failed to provide meaningful information about the constructs and had poor discriminating levels. Through a second look at item 7 ("*Others around me bully me*"), it was noted that besides the poor wording, the item may be uncomfortable for participants to respond to, and hence it was decided that the item be excluded in the subsequent round of data collection for CFA. Item 1 ("*I find it easy to make friends and be with people*") was eliminated due to similar reasons (discomfort in responding). Item 5 ("*I find it difficult to concentrate on my studies/work at hand*") was originally included to examine the effects of social and emotional skills on academic performance, even though it did not properly fall into the either category of social or emotional competence. The authors anticipate that the weird positioning of this item amongst other items related to social and emotional competence might have led to participants responding to it poorly. Consequently, it was decided that this item would be removed moving forward.

After deletion of items was confirmed, a second look at the eigenvalues brought attention to the fact that the eigenvalue for the third factor was very close to 1 (precisely, 1.07), hence it could be excluded. Consistent with results of the scree plot, a 2-factor solution was finalised. All items forming factor 1 pointed towards social and emotional tendencies while items forming factor 2 were theoretically the reverse - i.e., they indicated social and emotional challenges. Hence, factor 1 was named as "Social and Emotional Orientations" while factor 2 was called "Social and Emotional Challenges". As one would expect, there was a significant negative correlation between the two ($r = -.35$, $p < .01$).

Confirmatory Factor Analysis

Prior to conducting the CFA, univariate and multivariate normality were examined on the data collected in the second round ($N = 1508$). Since the assumption of normality was not met (all p-values < .001), robust ML estimation was performed. This method uses Satorra-Bentler correction (Satorra & Bentler, 1994) while fitting a CFA model. Using the *lavaan* package in R, a two-factor model for social and emotional orientations and social and emotional challenges was created. Items were specified to load onto their hypothesised factors, and their variances were fixed at unity in order to standardise them. In total, 17 parameters were estimated: 8 error variances, 8 paths and 1 correlation.

The CFA model in this study demonstrated satisfactory fit: $\chi^2_{S\text{-}B}/df = 3.26$; robust CFI = .966; robust RMSEA (90% C.I.) = 0.039 (.028–.050); robust SRMR = 0.052; robust GFI = 0.994. Additionally, a second order CFA model was specified to check the presence of a higher order latent factor. An unsatisfactory fit for this model confirmed the existence of only first order factors. All factor loadings were significant at $p < .01$ and ranged from .36 to .88. They have been represented graphically along with their 95% confidence estimates in Figure 4. The two factors were negatively correlated ($r = -.27$, $p < .01$).

Other Psychometric properties

Reliability. Reliability analysis was performed by examining the values of Cronbach's alpha. Alpha values for the CFA model were below recommended thresholds: $\alpha_{SEO} = 0.69$ and $\alpha_{SEC} = 0.57$ (must be >.60 based on the recommendations of Hair et al., 1995). However, the value of Cronbach's alpha (α) is dependent on the number of items in a factor, as well as the number of points on the response scale, according to the formulae given below (Cronbach, 1951)

$$\alpha = \frac{k}{k-1}\left[1 - \frac{\sum_{i=1}^{k} s_i^2}{s_X^2}\right] \tag{1}$$

where k is the number of items in the factor, s_i^2 is the variance of the individual item i where $i = 1,\ldots,k$, and s_X^2 is the variance of all items forming the factor. This formula is also often expressed as

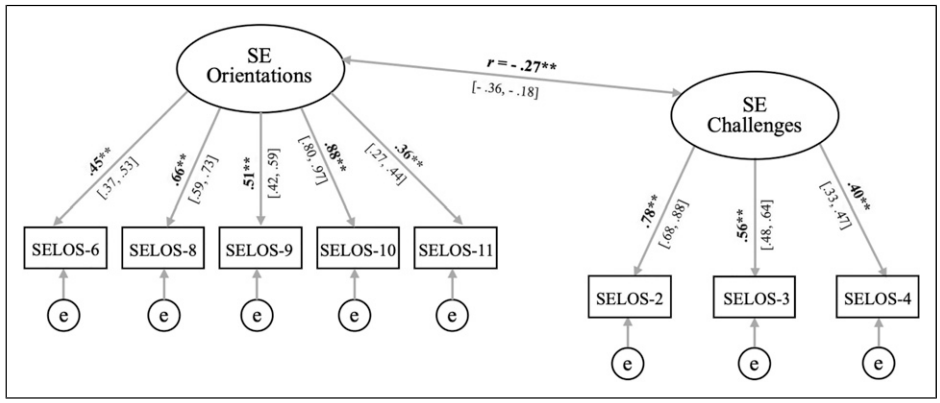

Figure 4. 2-Factor CFA model for the developed scale (**p < .01) - social and emotional orientations (SELOS items 6, 8, 9, 10, 11) and social and emotional challenges (SELOS items 2, 3 and 4).

$$\alpha = \frac{\underline{r}.k}{\left[1 + (k-1).\underline{r}\right]} \tag{2}$$

where k is the number of items in the factor and $\underline{r}$ is the average inter-item correlation. Therefore, a smaller number of items in the factor (k) reduces the alpha value. Additionally, a smaller number of points on the response Likert-scale (three, in our case) reduces the average inter-item correlation ($\underline{r}$), which consequently reduces the alpha value. The calculation of alpha is also based on the assumptions of tau equivalence and continuous nature of item responses. Due to such shortcomings, researchers have indicated that Cronbach's alpha has limited usefulness in terms of reporting reliability, especially for shorter scales (Sijtsma, 2008). A better alternative appears to be using McDonald's omega (Hayes & Coutts, 2020). Omega estimates based on polychoric correlations (scale items treated as polytomous/ordinal variables instead of continuous because they were responded to on a 3-point scale; Gadermann et al., 2012) revealed satisfactory values for the scale: $\omega_{SEO} = 0.71$, $\omega_{SEC} = 0.64$.

Discriminant Validity. In the absence of administering other measures in conjunction with the SELOS, discriminant validity of the scale was assessed through the CFA model itself by: 1) inspection of cross-loadings (Gefen & Straub, 2005), 2) the Fornell-Larcker criterion (Fornell & Larcker, 1981), and the more recent 3) Heterotrait-Monotrait ratio of correlations (HTMT; Henseler et al., 2014). Each item correlated weakly with the factor that it was not theoretically associated with, therefore, inspection of cross-loadings revealed sufficient discriminant validity. The AVE (Average Variance Extracted) for each factor ($AVE_{SEO} = 0.359$ and $AVE_{SEC} = 0.363$) was greater than the squared correlation between the factors (0.072), and hence the Fornell-Larcker criterion was satisfied. Finally, an HTMT value of 0.32 between the two factors indicated sufficient discriminant validity (HTMT value must be < 0.85 according to rules of Kline, 2011).

Measurement Invariance. A multifactorial invariance analysis was performed on the scale to establish its psychometric equivalence across males and females. This was done on the basis of the 4 steps outlined by Widaman and Reise (1997). The first step involves testing for configural invariance (i.e., whether basic organization of the factors is supported in the groups, which are the two genders in our case). When configural invariance is established, the next step tests for metric invariance (i.e., whether each item contributes to the latent factor to a similar degree across groups). It is tested by restricting factor loadings to be equivalent in the groups. The model with constrained factor loadings is then compared with the configural invariance model to determine fit. If the model fit is significantly worse compared to the configural invariance model, metric invariance is not supported. The next step is to test for scalar invariance (i.e., whether mean differences in the latent factor capture all mean differences in the shared variance of the items). It is tested by restricting the item intercepts to be equivalent in the groups, while retaining the constraints from the last step. This model is then compared with the model in the last step and fit is determined. Finally, residual invariance (i.e., whether the sum of specific variance and error variance is similar across groups) is tested by restricting the item residuals to be equal across the two groups. This model is again compared to the model in the last step and fit is determined (Putnick & Bornstein, 2016).

Based on the 4-step process laid out, starting from step 1, models were progressively restricted and their fit was compared with models in preceding steps using delta values for fit indices. The kind of fit indices that must be reported for these models is a source of debate, but most scholars recommend reporting the chi-square statistic ($\chi 2$) and 2–4 alternative fit indices such as the CFI, TLI, RMSEA, SRMR, etc. (Putnick & Bornstein, 2016). However, $\chi 2$ statistics are overly

sensitive to small, unimportant deviations from a "perfect" model in large samples and hence the focus in research has now shifted to using only alternative fit indices (French & Finch, 2006; Chen, 2007). Based on recommendations of Cheung and Rensvold (2002) and Chen (2007), CFI, RMSEA and SRMR were used to establish measurement invariance: ΔCFI < 0.01, ΔRMSEA < 0.015, and ΔSRMR < 0.030 (for metric invariance) or 0.015 (for scalar or residual invariance). The results of the multifactorial invariance analysis for SELOS (as seen from ΔCFI, ΔRMSEA and ΔSRMR values in Table 2) indicated that while all 4 invariance conditions were satisfied by ΔRMSEA and ΔSRMR values, residual and scalar invariance were not established from the ΔCFI values (values > 0.01). Several studies have shown that Cheung and Rensvold's (2002) cutoff for ΔCFI may not be appropriate in certain measurement design conditions, particularly in educational measurement (Jin, 2020; Khojasteh & Lo, 2015) or with dichotomous response scales. These cutoffs also work best for equal sample sizes in general (Chen, 2007). Further detailed research would need to be conducted to establish complete measurement invariance of the SELOS keeping in mind these factors.

Gender and Age Differences

After establishing the factor structure, factor scores were examined for differences in gender and age. Univariate normality of scores was examined using visual inspection of Q-Q plots. Homogeneity of variance was assessed using Levene's Test for equality of variances (Levene, 1960). Wherever the distribution was heteroscedastic, Welch's statistics were reported. Effect size measure, Cohen's *d* was calculated and interpreted according to rules of Cohen (1988): 0.20 = small, 0.50 = medium, and 0.80 = large. An unpaired *t*-test revealed that females ($M = 1.74$, $SD = 0.38$) scored significantly higher than males ($M = 1.67$, $SD = 0.43$) on the subscale of social and emotional challenges ($t(1167.32) = 3.25$, $p < .01$, $d = 0.17$). Additionally, even though females ($M = 2.77$, $SD = 0.28$) scored higher than males ($M = 2.74$, $SD = 0.30$) on social and emotional orientations, the *t*-test revealed an insignificant result ($p > .05$).

The influence of age on factor scores was examined using regression modelling. An insignificant effect of age ($p > .05$) was observed on social and emotional orientation scores. However, a significant effect of age ($R = 0.083$, $p < .01$) was observed for social and emotional challenges, although weak in magnitude.

Discussion

Most measures of SEL that exist currently have been developed in English and with Western populations. Using these measures with different cultural populations poses several challenges during translation and adaptation. Moreover, the process of translation does not ensure cultural

Table 2. Multifactorial invariance analysis of the scale by gender (Model 1 = configural invariance, Model 2 = metric invariance, Model 3 = scalar invariance, and Model 4 = residual invariance).

Model	CFI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
Model 1	0.964	0.041	0.061	-	-	-
Model 2	0.955	0.042	0.065	0.009	0.001	0.004
Model 3	0.938	0.047	0.066	0.018	0.005	0.001
Model 4	0.904	0.057	0.066	0.034	0.010	0.000

Note. Models test the two-factor model, estimated using robust ML estimation. As $\chi 2$ statistics are affected by sample size, invariance decisions were made based on the CFI, RMSEA and SRMR.

transferability of measures that have been designed keeping in mind Western audiences. There is a paucity of culturally sensitive measures of SEL in languages other than English. The primary aim of this study was to create a brief, psychometrically validated scale in Hindi for measuring social and emotional learning skills of students in India. A total of 11 items measuring social and emotional competence were created based on review of items from the GSHS and delivered to government school students in the Indian state of Rajasthan. Findings indicated the presence of a two-factor model comprising social and emotional orientations and social and emotional challenges. This factor structure was confirmed by satisfactory fit obtained in the CFA. Additionally, the developed scale demonstrated good reliability, discriminant validity and partial measurement invariance across the two genders. The final 8-item scale is presented in Appendix A.

Although it seems important to investigate the level of satisfaction of SEL surveys, namely do students respond carelessly to questions, the current scale made every attempt to use simple language and avoid embarrassing questions, to enable students across a wide range of reading levels and grades to respond to it. Additionally, all possible efforts were made to eliminate the bias of social desirability by collecting data anonymously, thereby ensuring that the responders marked responses honestly.

We also found that females reported higher socio-emotional challenges as compared to males. It is unclear whether both females and males report emotional states that are expected of them rather than what they experience. It is possible that this finding reflects the prevalent gender norms that are emphasized in society where females are expected to display more docile character traits while males are expected to display more aggressive masculine behaviors (Koenig, 2018).

Limitations

There were several limitations to this study. While our sample was large, it was not representative of the Indian student population as a whole. Testing similar questionnaires in other Indian languages is likely to present a more holistic picture of the SEL states of students in the country. Convergent validity of the scale could not be examined since there are no similar measures of SEL available in Hindi to the best of our knowledge. There are a few other related psychometrically validated scales that could have been delivered to assess convergent/predictive validity such as the Hindi versions of the Depression, Anxiety, and Stress Scale - 21 (DASS-21; Singh et al., 2015), Strengths and Difficulties Questionnaire (SDQ; Singh et al., 2015), and the Self Efficacy Scale (SES; Shrivastava et al., 2018). However, these were avoided to maintain brevity of the assessment, as well as avoid embarrassing/stigmatizing questions because they are translated versions of the English measures and have not been developed locally. Having said that, it is recognized that establishing convergent and predictive validity of the scale is important for further research through use of more appropriate scales or collection of external variables through interviews and qualitative surveys (this would have been possible if the study was not conducted remotely). Another important psychometric property of the scale, test-retest reliability, could not be examined even though the participants retook the scale after a few months, because the development and validation of the SELOS was a part of a larger study meant to evaluate the efficacy of an SEL intervention in Hindi. Since the two points of time at which the participants took the scale were not independent, rather consisted of new learning (an intervention) between them (the scale was used as a pre-post measure), it would not be fair to compare their scores for test-retest reliability. A spin-off research could also be undertaken from this study by asking participants to respond to the items of SELOS on both 3-point and 5-point response scales and examining how psychometric properties differ under the two cases.

The authors recognize that the purpose of this study was to create a brief scale that can be used quickly with a large audience, however, we do also understand the need to develop larger, more

nuanced versions of such scales to gain complete understanding of SEL states in students in India. We hope that the SELOS can serve as a starting point for psychometrically-validated SEL measurement in India. Extending this research to develop translated versions of the questionnaire in major Indian languages such as Marathi, Oriya, Bengali, Telugu, etc. (since a considerable population of India does not speak Hindi) will make it possible to conduct a nationwide assessment of SEL skills in India. Since the SELOS is simple, easy to administer and can be delivered online, we hope it will be widely used by both governments and schools globally.

Appendix A

Validated Scale (SELOS)

(Question)	कभी नहीं (Never)	कभी कभी (Sometimes)	हमेशा (Always)
1. मैं अकसर उदास और नाखुश रहता/रहती हूँ। (I experience sadness and unhappiness often)	☐	☐	☐
2. मुझे अकसर गुससा आता है। (I often get angry and lose my temper)	☐	☐	☐
3. मैं अपने हमउम् साथियों के साथ रहने की बजाये अकेले रहना पसंद करता/करती हूँ। (I prefer to be alone rather than with people my age)	☐	☐	☐
4. मैं कुछ करने से पहले सोचता/सोचती हूँ। (I think before I do things)	☐	☐	☐
5. मैं सदैव दूसरों की भावनाओं का धयान रखता/रखती हूँ। (I pay attention to the feelings of others)	☐	☐	☐
6. मैं अपने शिक्षकों से सहजता से बात कर सकता/सकती हूँ। (I talk easily with teachers)	☐	☐	☐
7. मुझे दूसरों की मदद करना अच्छा लगता है। (I like to help others)	☐	☐	☐
8. मुझे लगता है कि सीखने में भावनाएँ मायने रखती हैं। (I think feelings matter in learning)	☐	☐	☐

Note. Items 1, 2 and 3 form the subscale of social and emotional challenges, and items 4, 5, 6, 7, and 8 form the subscale of social and emotional orientations

ORCID iDs

Mayank Sharma https://orcid.org/0000-0003-1119-4675
Sonal Chheda https://orcid.org/0000-0002-7352-8223
Reshma Piramal https://orcid.org/0000-0003-0609-5242
Neha Bhatia https://orcid.org/0000-0003-4043-5159
Tyralynn Frazier https://orcid.org/0000-0002-7450-0556
Nandini Chatterjee Singh https://orcid.org/0000-0002-0005-5541

Funding

This research was funded by intramural funding from the UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development, New Delhi.

Declaration of Conflicting Interests

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

Ethics Declaration

The entire research study was carried out in strict adherence to ethical research practices, and in accordance with the principles of UNESCO's Universal Declaration on Bioethics and Human Rights (available at https://unesdoc.unesco.org/ark:/48223/pf0000146180).

References

Anthony, C. J., Brann, K. L., Elliott, S. N., & Garis, E. J. (2021). *Examining the structural validity of the SSIS SEL brief scales—teacher and student forms. Psychology in the Schools*. Wiley.

Baker, F. B (2001). *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation.

Banerjee, R., Weare, K., & Farr, W. (2013). Working with 'Social and Emotional Aspects of Learning' (SEAL): Associations with school ethos, pupil social experiences, attendance, and attainment. *British Educational Research Journal*, *40*(4), 718–742. https://doi.org/10.1002/berj.3114

Bartlett, M. S. (1954). A Note on the multiplying factors for various $\chi 2$ approximations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *16*(2), 296–298. https://doi.org/10.1111/j.2517-6161.1954.tb00174.x

Bentler, PM (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Boncu, A., Costea, I., Costea, I., & Minulescu, M. (2017). A meta-analytic study investigating the efficiency of socio-emotional learning programs on the development of children and adolescents. *Romanian Journal of Applied Psychology. 19*(2), 35–41. https://doi.org/10.24913/rjap.19.2.02

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005.

Cattell, R. B (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Census of India (2018). *Abstract Of Speakers' Strength Of Languages And Mother Tongues - 2011*. Retrieved 27 July, 2021. https://censusindia.gov.in/2011Census/Language-2011/Statement-1.pdf

Cha, E.-S., Kim, K. H., & Erlen, J. A (2007). Translation of scales in cross-cultural research: Issues and techniques. *Journal of Advanced Nursing*, *58*(4), 386–395. https://doi.org/10.1111/j.1365-2648.2007.04242.x

Chambers, C. T., & Johnston, C (2002). Developmental differences in children's use of rating scales. *Journal of pediatric psychology. 27*(1). 27–36. https://doi.org/10.1093/jpepsy/27.1.27

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*. *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*. *9*(2), 233–255. https://doi.org/10.1207/s15328007sem0902_5

Cohen, J (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers.

Credé, M, Harms, P, Niehorster, S, & Gaye-Valentine, A (2012). An evaluation of the consequences of using short measures of the big five personality traits American psychological association (APA). *Journal of Personality and Social Psychology*, *102*(4), 874–888. https://doi.org/10.1037/a0027403

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests springer ccience and business media LLC. *Psychometrika*. *16*(3), 297–334. https://doi.org/10.1007/bf02310555

Denham, S. A., Ferrier, D. E., Howarth, G. Z., Herndon, K. J., & Bassett, H. H. (2016). Key considerations in assessing young children's emotional competence. *Cambridge Journal of Education*, *46*(3), 299–317. https://doi.org/10.1080/0305764x.2016.1146659

Department of Education, State Government of Rajasthan (2018). *School education: State report card 2018-19*. Retrieved 28 October, 2021. https://education.rajasthan.gov.in/content/dam/doitassets/education/school%26secondaryeducation/rcee-jaipur/pdf2017/PDF2019/REPORT%20CARD%202018-19.pdf

Devellis, R (2012). *Scale development theory and applications*. Sage Publications.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development. Wiley. 82*(1). 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Edwards, P., Roberts, I., Sandercock, P., & Frost, C. (2004). Follow-up by mail in clinical trials controlled clinical trials: Does questionnaire length matter? *Elsevier BV. 25*(1). 31–52. https://doi.org/10.1016/j.cct.2003.08.013

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M T., Haynes, N. M., Kessler, R., Schwab–Stone, M. E., & Shriver, T. P (1997). *Promoting social and emotional learning: Guidelines for educators*. Association for Supervision and Curriculum Development.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299. https://doi.org/10.1037/1082-989x.4.3.272

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39. https://doi.org/10.2307/3151312

French, B. F. & Finch, W. H. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *13*(3), 378–402. https://doi.org/10.1207/s15328007sem1303_3

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *University of Massachusetts Amherst*. *17*(3). https://doi.org/10.7275/N560-J767

Gefen, D., & Straub, D (2005). A practical guide to factorial validity using PLS-graph: Tutorial and annotated example association for information systems. *Communications of the Association for Information Systems*, *16*(1). https://doi.org/10.17705/1cais.01605

Gorsuch, R (1983). *Factor analysis* (2nd ed.). Lawrence Erlbaum Associates.

Government of India (2018). *Census of India 2011: Language, India, States and Union Territories*. Retrieved 25 July, 2021. https://censusindia.gov.in/2011Census/C-16_25062018_NEW.pdf

Gururaj, G., Varghese, M., Benegal, V., Rao, G.N., Pathak, K., Singh, L.K., Mehta, R.Y., Ram, D., Shibukumar, T.M., Kokane, A, Lenin Singh, R.K., Chavan, B.S., Sharma, P., Ramasubramanian, C., Dalal, P.K., Saha, P.K., Deuri, S.P., Giri, A.K., Kavishvar, A.B., & Misra, R., NMHS collaborators group (2016). National mental health survey of india, 2015-16: Summary. *National Institute of Mental Health and Neuro Sciences, NIMHANS Publication No*, *59*(1). https://doi.org/10.4103/psychiatry.IndianJPsychiatry_102_17

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C (1995). *Multivariate data analysis*. Prentice-Hall.

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than cronbach's alpha for estimating reliability. But. *Communication cMethods and Measures*, *14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling springer science and business media LLC. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Henze, N., & Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Communications in Statistics - Theory and Methods*, *19*(10), 3595–3617. https://doi.org/10.1080/03610929008830400

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 17–185. https://doi.org/10.1007/BF02289447

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and Psychological Measurement*, *71*(4), 617–637. https://doi.org/10.1177/0013164410382896

Jin, Y. (2020). A note on the cutoff values of alternative fit indices to evaluate measurement invariance for ESEM models. *International Journal of Behavioral Development*. *44*(2) 166–174. https://doi.org/10.1177/0165025419866911

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/bf02291575

Kemper, C. J., Trapp, S., Kathmann, N., Samuel, D. B., & Ziegler, M. (2018). Short versus long scales in clinical assessment: Exploring the trade-off between resources saved and psychometric quality lost using two measures of obsessive–compulsive symptoms. *SAGE Publications*. *26*(5), 767–782. https://doi.org/10.1177/1073191118810057

Khojasteh, J., & Lo, W.-J. (2015). Investigating the sensitivity of goodness-of-fit indices to detect measurement invariance in a bifactor model. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(4), 531–541. https://doi.org/10.1080/10705511.2014.937791

Kline, R. B (2011). *Principles and practice of structural equation modeling*. (3rd ed.). Guilford Press.

Koenig, AM (2018). Comparing prescriptive and descriptive gender stereotypes about children, adults, and the elderly. *Frontiers in Psychology*, *9*, 1086. https://doi.org/10.3389/fpsyg.2018.01086

Kruyen, P. M., Emons, W. H. M., & Sijtsma, K. (2013). On the shortcomings of shortened tests: A literature review. *International Journal of Testing*, *13*(3), 223–248. https://doi.org/10.1080/15305058.2012.703734

Lagi, R., & Armstrong, D (2017). The integration of social and emotional learning and traditional knowledge approaches to learning and education in the pacific, In E. Frydenberg, A. J. Martin, & R. J. Collie (Eds.), *Social and emotional learning in Australia and the Asia-Pacific*. Springer.

Levene, H (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to Probability and Statistics* (pp. 278–292). Stanford University Press.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99. https://doi.org/10.1037/1082-989x.4.1.84

Masten, A. S (2001). Ordinary magic: Resilience processes in development. *American Psychologist*, *56*(3), 227–238. https://doi.org/10.1037/0003-066x.56.3.227

Mellor, D., & Moore, K. A. (2013). The use of likert scales with children. *Journal of pediatric psychology*. *39*(3). 369–379. https://doi.org/10.1093/jpepsy/jst079

Merrell, K. W (2002). *School social behavior scales* (Second Edition). Assessment-Intervention Resources.

Ministry of Human Resource Development, Government of India (2020). *National education policy 2020*. Retrieved 27, July. https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf

Moore, K. A., & Mellor, D. J. (2003). The nature of children's social interactions at school. School psychology international. *SAGE Publications*. *24*(3), 329–339. https://doi.org/10.1177/01430343030243005

Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S., & Ferreira, M. E. C (2017). Ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica*. *30*(1). https://doi.org/10.1186/s41155-016-0057-1

National Council for Educational Research and Training. (2005). *National Curriculum Framework 2005*. Retrieved 18 March, 2022. https://ncert.nic.in/pdf/nc-framework/nf2005-english.pdf

Nese, R. N. T., Doerner, E., Romer, N., Kaye, N. C., Merrell, K. W., & Tom, K. M. (2012). Social emotional assets and resilience scales: Development of a strength-based short-form behavior rating scale system. *Journal for Educational Research Online*. *4*(1), 124–139. Waxmann. https://doi.org/10.25656/01

Oberle, E., Domitrovich, C. E., Meyers, D. C., & Weissberg, R. P. (2016). Establishing systemic social and emotional learning approaches in schools: a framework for schoolwide implementation. *Cambridge Journal of Education*, *46*(3), 277–297. https://doi.org/10.1080/0305764x.2015.1125450

Oishi, S (2007).The application of structural equation modeling and item response theory to cross-cultural positive psychology research. In A. D. Ong & M. H. M. van Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 126–138). Oxford University Press.

Payton, J., Weissberg, R.P., Durlak, J.A., Dymnicki, A.B., Taylor, R.D., Schellinger, K.B., & Pachan, M (2008). *The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews*. Collaborative for Academic, Social, and Emotional Learning.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*. *41*. 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Reynolds, C.R., & Kamphaus, R.W (2015). *Behavior assessment system for children [Assessment instrument]* (3rd ed.). Pearson.

Rimm-Kaufman, S. E., Larsen, R. A. A., Baroody, A. E., Curby, T. W., Ko, M., Thomas, J. B., Merritt, E. G., Abry, T., & DeCoster, J. (2014). Efficacy of the responsive classroom approach. *American Educational Research Journal*, *51*(3), 567–603. https://doi.org/10.3102/0002831214523821

Satorra, A., & Bentler, P. M (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research*. (pp. 399–419). Sage.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, *8*(2), 23–74

Schonert-Reichl, K. A., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T. F., & Diamond, A (2015). Enhancing cognitive and social–emotional development through a simple-to-administer mindfulness-based school program for elementary school children: A randomized controlled trial. *Developmental Psychology*, *51*(1), 5–66. https://doi.org/10.1037/a0038454

Schumacker, R. E., & Lomax, R. G (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Lawrence Erlbaum

Shapiro, S. S., & Wilk, M. B (1965). An analysis of variance test for normality (Complete Samples). *Biometrika*, *52*(3/4), 591. https://doi.org/10.2307/2333709

Shrivastava, P., Sonekar, L., & Janghel, G (2018). Development of psychometric properties in self-efficacy scale in Hindi language (SE scale in Hindi version). *International Journal of Indian Psychology*, *6*(2), 50–55. https://doi.org/10.25215/0602.086

Sijtsma, K (2008). On the use, the misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/s11336-008-9101-0

Singh, K, Junnarkar, M, & Sharma, S (2015). Anxiety, stress, depression, and psychosocial functioning of Indian adolescents. *Indian Journal of Psychiatry*, *57*(4), 367–374. https://doi.org/10.4103/0019-5545.171841

Tabachnick, B.G., & Fidell, L.S (2001). *Using multivariate statistics*. 4th edition. Allyn & Bacon.

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*. *88*(4), 1156–1171. https://doi.org/10.1111/cdev.12864

Torrente, C., Alimchandani, A., & Aber, J. L (2015). International perspectives on SEL. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 566–587). Guilford.

Unified District Information System for Education (UIDSE).(2019). *Statistics Dashboard*. Retrieved 27, July, 2021. https://dashboard.udiseplus.gov.in/#/reportDashboard/sReport

United Nations Children's Fund (UNICEF) (2020). *General country profile – India*. Retrieved 25 July, 2021. https://www.unicef.org/rosa/media/10566/file/India.pdf

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327. https://doi.org/10.1007/BF02293557

Velicer, W.F., Eaton, C.A., & Fava, J.L (2000).Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R.D. Goffin & E Helmes (Eds.), *Problems and Solutions in Human Assessment*. Springer.

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, *25*(1), 1–28. https://doi.org/10.1207/s15327906mbr2501_1

Wang, WL, Lee, HL, & Fetzer, SJ (2006). Challenges and strategies of instrument translation. *Western Journal of Nursing Research*, *28*(3), 310–321. https://doi.org/10.1177/0193945905284712

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). American Psychological Association. https://doi.org/10.1037/10222-009

World Health Organization (WHO) (2021). *Global school-based student health survey (GSHS)*. Retrieved 27 July, 2021. http://www.who.int/chp/gshs/en/

World Health Organization (WHO) (2017). *Mental health status of adolescents in South-East Asia: Evidence for action*. World Health Organization. https://apps.who.int/iris/handle/10665/254982

Wright, KD, & Asmundson, GJ (2003). Health anxiety in children: Development and psychometric properties of the childhood illness attitude scales. *Cognitive Behaviour Therapy*, *32*(4), 194–202. https://doi.org/10.1080/16506070310014691

Zins, J. E., & Elias, M. J (2007). Social and emotional learning: Promoting the development of all students. *Journal of Educational and Psychological Consultation*, *17*(2–3), 233–255. https://doi.org/10.1080/10474410701413152