frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 01 July 2025
DOI 10.3389/feduc.2025.1534413

OPEN ACCESS

EDITED BY
Yong Luo,
Pearson, United States

REVIEWED BY
Qianqian Pan,
Nanyang Technological University, Singapore
Yanyan Tan,
Pearson, United States

*CORRESPONDENCE
Jorge E. Gonzalez
jgonza14@trinity.edu

†These authors have contributed equally to this work and share first authorship

RECEIVED 25 November 2024
ACCEPTED 14 May 2025
PUBLISHED 01 July 2025

CITATION
Kim H and Gonzalez JE (2025) Exploring the dimensionality of the CLASS Pre-K in Latine preschool classrooms: further support for a bifactor model.
Front. Educ. 10:1534413.
doi: 10.3389/feduc.2025.1534413

Exploring the dimensionality of the CLASS Pre-K in Latine preschool classrooms: further support for a bifactor model

Hanjoe Kim[1†] and Jorge E. Gonzalez[2*†]

[1]Department of Psychology, Yonsei University, Seoul, Republic of Korea, [2]Department of Education, Trinity University, San Antonio, TX, United States

Introduction: The Classroom Assessment Scoring System (CLASS) is a classroom observational tool commonly used to evaluate both structural and process elements of a classroom. Despite its widespread use, over the last decade, debates about its factor structure across various cultural contexts or settings have arisen. In this study, we asked the question, does the CLASS Pre-K retain its established three-domain structure when applied to classrooms with both Latine teachers and preschool children as compared to its original validation in broader U.S. populations? Six alternative models that were suggested by literature were explored.

Methods: Using a series of confirmatory factor analyses, we found that a best fitting bifactor model with one generic factor (Responsive Teaching) and three correlating specific factors (Emotional Support, Classroom Organization, Instructional Support) aligned with the developer's theorizing was a good fit for the data.

Results: The generic factor represents a common dyadic systems-level property across all CLASS Pre-K dimensions, which can be characterized by the construct of responsivity.

Discussion: Furthermore, Latine cultural values such as *bien educado* and *educación* are illustrated as a plausible explanation of the factor structure found in this study. Caution should be exercised in using the CLASS Pre-K in classrooms with cultural backgrounds that differ from the original norming sample, as its domains may not fully capture or accurately reflect culturally specific teacher-child interactions.

KEYWORDS

classroom, observation, cultural, factor analysis, Latine

Introduction

Decades of research has documented that attending preschool benefits children's language development and socioemotional learning, especially when these programs are of high quality (Karlsen et al., 2024). That is, high-quality early childhood education (ECE) programs promote key developmental skills, creating a strong foundation for future academic and social success (Slot et al., 2018). Further, it is also well established that the quality of ECE programs is multidimensional encompassing both structural and process elements (Bihler et al., 2018).

Among the instruments developed to assess the structural and process quality of ECE programs is the Classroom Assessment Scoring System Pre-K or CLASS (Pianta et al., 2008). The CLASS is widely used in the United States and abroad; however, recent international

findings have raised questions about its dimensionality. Specifically, despite its widespread use, debates about its factor structure across various cultural contexts or settings have arisen over the last decade (Ng et al., 2021). This study contributes to existing factor structure debates by focusing on the dimensionality of the CLASS Pre-K from observations in primarily Latine preschool classrooms taught by Latine teachers.

Young children's early educational experiences are foundational in promoting academic learning success. Extensive evidence from longitudinal studies in both developed and developing countries has consistently demonstrated that quality ECE produces significant long-term benefits, extending to academic achievement and overall life success (McDoniel et al., 2022). As such, policymakers, researchers, and practitioners, both nationally and internationally, underscore the critical importance of delivering high-quality ECE as a protective mechanism, particularly for children from less advantaged backgrounds. There remains, however, an urgent need to identify, evaluate, and, whenever necessary, implement changes in classroom-level processes predictive of these long-term benefits (Brown et al., 2010).

Class

To that end, the development of standardized observational measures of classroom quality has significantly advanced the field of ECE by enabling the systematic identification and tracking of goals for improvement (McDoniel et al., 2022). One such observational instrument, CLASS, is the most widely used classroom observational system in practice, policy, and research (Bihler et al., 2018; Gordon and Peng, 2020). Developed based on developmental theory and empirical evidence, the CLASS emphasizes teacher-student interactions as the primary catalyst for children's learning and development (Gordon and Peng, 2020). The CLASS employs a structured observational protocol to assess three key teacher-child interaction domains: Emotional Support, Classroom Organization, and Instructional Support. Each domain is further subdivided into three to four specific dimensions that are evaluated based on underlying behavioral indicators observed in the classroom setting (Pianta et al., 2008). Emotional Support evaluates the extent to which teachers foster a positive classroom climate through their daily interactions with students; Classroom Organization examines the effectiveness of routines and procedures in managing children's behavior, time, and attention within the classroom; and Instructional Support assesses the strategies teachers use to implement the curriculum and promote cognitive and language development (Pianta et al., 2008).

The CLASS's conceptual foundation builds on the "teaching through interaction model" developmental framework, which proposes that student learning and developmental gains are driven by teacher-student interactions (Hamre et al., 2013). As a result, this framework suggests that a nuanced and detailed assessment of the quality of teacher-child interactions can offer critical insights into their impact on child development and educational outcomes (Ng et al., 2021). Indeed, the quality of teachers' emotional and organizational support over time plays a significant role in the development of children's competencies, particularly socio-emotional skills, literacy and language skills, the management of problematic behaviors, and relationships with peers, often irrespective of a family's income, ethnicity, or maternal education level (Bihler et al., 2018; Iruka et al., 2022).

Theoretically, the teaching through interactions framework layers within a broader developmental-ecological framework (Bronfenbrenner and Morris, 1988), which provides a robust understanding of the dyadic teacher-child interactions as sources of child social, cognitive, and behavioral development (Bronfenbrenner and Morris, 1988). In this model, therefore, it is the proximal processes or the interactions between teachers, peers, and instructional materials, that are the primary drivers of children's learning (Brown et al., 2010).

Ecocultural theory

Another framework useful in contextualizing the teaching through interactions model is ecocultural theory. Ecocultural theory extends Bronfenbrenner's model by highlighting that variation in children's behavior is in part a function of cultural boundedness in terms of class, heritage, and language and expressed through reciprocal interactions (e.g., teacher-student) in the immediate environment such as a classroom (Bridges et al., 2012). As such, ecocultural theory captures the nuanced interplay between individual-level processes and the broader socio-contextual conditions that shape them. From an ecocultural theory perspective, therefore, daily routines and activities are structured in accordance with cultural values, available economic and social resources, community norms, and specific contextual demands. These routines, particularly those that promote children's learning (e.g., teacher-child instructional routines), are hypothesized to mediate developmental outcomes within cultural niches (e.g., families, classrooms) (McWayne et al., 2016).

Factor analysis and the theorized structure of CLASS

The theorized structure of the CLASS can be validated and examined using factor analysis. That is, factor analysis allows assessing whether the developer-proposed three-factor structure provides a better fit than alternative structures, such as bi- or single-factor models. Factor analysis can also determine if the three factors are distinct, with small-to-moderate intercorrelations, and whether items load appropriately onto the proposed domains (Gordon and Peng, 2020).

Regarding the CLASS Pre-K, few factor analytic studies have been published, and existing studies provide mixed support for the proposed structure, with some finding support for alternative models. Preliminary research on the CLASS Pre-K has demonstrated that the proposed three-factor structure is more effective than a one- or two-domain factor structure although it demonstrated suboptimal fit indices (Hamre et al., 2013). However, recently there has been debate concerning the underlying factorial structure of the framework. Specifically, some studies have found the three-factor model to have inadequate fit, even after post-hoc modifications (Bihler et al., 2018). Furthermore, several studies have allowed for correlated residuals, indicating that the dimensions underlying the CLASS may share a

common cause that is not accounted for by the underlying "teaching through interactions" model, suggesting that an unknown factor may be influencing teacher-child interactions.

In response to these concerns, Hamre et al. (2014) put forth a modified bifactor model that proposed two domains, one general domain-factor (responsive teaching) and two domain-specific factors (Positive Management and Routines, and Cognitive Facilitation), specified to be uncorrelated with each other. This model provides more clarity due to the completely uncorrelated nature of the three factors (Hamre et al., 2014), and by offering an alternative structural representation of multidimensionality, bifactor measurement models have been proposed as a means of addressing problems in conceptualizing and measuring psychological constructs (Ng et al., 2021).

Across fields, bifactor model structures have become popular due to their potential to account for high correlations between and among conceptually distinct constructs (Gordon and Peng, 2020). However, some researchers caution that bifactor models are limited due to factor convergence difficulties, large samples requirements, and often being based on empirical rather than theoretical foundations (Li et al., 2020). For example, a test of the bifactor structure by Hamre et al. (2014) revealed only adequate fit among some of the fit indices with not all items loading meaningfully on a domain-specific factor along with convergence and identification problems with the bifactor structures. Gordon and Peng (2020), in turn, found that a revised three-domain factor structure based on substantive interpretation fitted better than the original purported three-factor structure. Specifically, in the revised three-domain structure, items in the domains of Emotional Support and Classroom Organization were reorganized into newly named domains: a Climate and Management domain and a Sensitivity and Regard domain while the Instructional Support domain remained intact.

From yet another perspective, given that classroom quality may be a culturally mediated construct, debates have arisen regarding the appropriateness of applying Western standards of quality in international and other contexts. Specifically, researchers have noted that the quality of teacher-child interactions is likely influenced by cultural concepts and values, which may vary significantly across different cultural settings. These values influence expectations around communication styles, respect for authority, collaborative vs. individual learning preferences, and emotional expressiveness, all of which play critical roles in shaping the dynamics of teacher-child interactions (Pastori and Pagani, 2017).

Hu et al. (2016), in examining the applicability of the CLASS in Chinese kindergarten classrooms, found support for both a three-factor and a bifactor model of teacher-child interaction quality, noting that some of the factor loadings aligned with known cultural expectations and practices between countries. In another study using German preschool classrooms, Bihler et al. (2018) also found support for the purported three-factor model, noting that the model demonstrated the best model to fit German preschool data; however, there was no support for the bifactor model. Finally, using data from early childhood centers in Norway, Karlsen et al. (2024) also found support for the three-factor structure but noted that the model required post-hoc adjustments to achieve adequate model fit.

In contrast, other researchers have reported the need for alternative models to the purported three-factor structure to ensure data fit (Ng et al., 2021). Using the CLASS Pre-K in Singapore preschool classrooms, Ng et al. (2021) found that the factorial structure of the CLASS was best conceptualized as bifactor comprising of the purported Responsive Teaching general factor and two specific factors: (a) Proactive Management and (b) Routines and Cognitive Facilitation. Nevertheless, these researchers identified some differences between their findings and those of Hamre et al. (2014), especially in terms of factor loadings for Proactive Management and Routines, with their study demonstrating that the strongest loadings were Negative Climate, Behavior Management, and Teacher Sensitivity whereas in Hamre et al. (2014) and Hu et al. (2016) the strongest loadings were Negative Climate, Behavior Management, and Productivity.

In summary, previous CFA studies both in the United States and abroad suggest that the three-factor model posited by the "teaching through interactions" developmental theoretical framework proposed by Hamre et al. (2014) fits the data across studies more robustly than either a one-, two-, or bifactor models. These findings along with those of other studies (Ng et al., 2021) suggest that an alternative bifactor model consisting of a general factor (i.e., Responsive Teaching) and two specific factors (i.e., Proactive Management and Routines, and Cognitive Facilitation) fit the data more robustly. Given these inconsistent findings, more studies are needed to investigate the underlying factor structure of the CLASS rather than assuming a three-factor structure, especially across sociocultural contexts or classroom settings. To begin to fill this gap in the literature, the present study explored the factorial structure of the CLASS Pre-K in preschool classrooms serving Latine preschoolers.

Current study

This study contributes to current debates about the factor structure and cultural applicability of the CLASS Pre-K across educational settings by focusing on preschool classrooms with Latine preschoolers taught by Latine teachers. It is altogether possible that differences in the sociocultural context between the setting in this study and others could yield a different internal factorial structure.

Method

Data for the study were collected as part of a federally funded study to investigate the effects of a content-rich interactive preschool shared-book reading program on children's receptive and expressive vocabulary. Classroom observations were collected in the fall (October) across three school years in classrooms serving dual language learners (DLL) taught by Latine Spanish- and English-speaking certified preschool teachers. Two school districts participated in the study.

Setting

The sample consisted of 137 DLL preschool classrooms across two school districts located in a state along the United States and Mexico border. Demographically, one of the school districts had a student

population of 24,040 students of which 99.8% were Latine, 93.2% economically disadvantaged, and 58% limited English proficient (LEP). Of the teachers, 97% were Latine; the student-teacher ratio was 13:4. The other school district had a student population of 29,242 students of which 99% were Latine, 92% economically disadvantaged, and 43% LEP. Of the teachers, 95% were Latine; the student teacher ratio was 13:7.

Sample

A total of 831 preschoolers across 137 preschool classrooms participated. Preschoolers were 98% Latine, 49% were female, and 97% were receiving free or reduced-price lunch. Teachers across the 137 classrooms were 95% Latine, with 97% being female. Most of the preschool classrooms had a bilingual paraprofessional teacher's aide in addition to the classroom teacher. All teachers held early childhood certifications. No demographic information was available on the paraprofessionals.

Procedures

The study was approved by the Institutional Review Board, and both the children's parents and their teachers consented in their preferred language (i.e., Spanish or English). Prior to collecting data, graduate student observers underwent training by a licensed CLASS trainer on adhering to standard procedures, subsequently passing a reliability test with a preset criterion of at least 80% agreement within one scale point. To obtain scores most typical of each classroom, observations were conducted in October of the school year to avoid beginning-of-the-year or holiday disruptions of standard classroom routines. To mitigate potential confounding effects associated with variation in teacher-child verbal interactions attributable to the language of instruction (i.e., Spanish versus English), all observations were scheduled and conducted during periods of Spanish language instruction. This decision was made because Spanish was the primary language spoken at home by the Latine children participating in the study. Additionally, both teachers and observers were fluent in Spanish and English, thus maintaining consistency in language use during data collection. To minimize observer bias, the trained CLASS Pre-K graduate student observers had never interacted with the teacher prior to the observation or been in the teacher's classroom. Observers rated classrooms along a 7-point scale from "low" to "high" with 1 and 2 being low range, 3, 4, and 5 being mid-range, and 6 and 7 classified being high range. To ensure robustness of the data, 20% of the observations were conducted by dual observers and double-coded, resulting in a high interrater reliability (89%).

Data analysis

A series of confirmatory factor models were estimated to examine the latent structures of a teacher-student interactions measure – CLASS (CLASS K-3) (Pianta et al., 2008). In addition, our study intended to validate the established three-factor 10-dimension structure reported for the CLASS instrument (see Pianta et al., 2008) in the context of Pre-K classrooms for ELLs. First, we estimated a three-factor CFA model (i.e., the original CLASS K-3 three-factor model) (see Pianta et al., 2008; Hamre et al., 2007) to evaluate Pre-K classroom observations. The three factors/constructs were: Emotional Support, Classroom Organization, and Instructional Support. Each construct was viewed as a major domain of the global measure of classroom quality.

Next, a two-factor CFA model was tested, which hypothesized that the 10 dimensions manifested only two domains: Emotional Support-Classroom Organization and Instructional Support. Last, a single-factor model was proposed with all 10 dimensions loaded on a single factor. In other words, the teacher-student interactions were measured as an overall classroom quality. In line with the recent literature (e.g., Sandilos and DiPerna, 2014), three additional models were also tested, including Sandilos and DiPerna's (2014) revised CLASS K-3 model, Pakarinen et al.'s (2010) revised CLASS K-3 model, and Hamre et al.'s (2014) CLASS K-3 bifactor model.

The six competing models were estimated using *Mplus* (Version 8) (Muthén and Muthén, 1998–2017) with the maximum likelihood estimation method. The following indices were used to evaluate the goodness-of-fit of each model: chi-square test statistics, root mean square error of approximation (RMSEA), comparative fit index (CFI), and standardized root mean square residual (SRMR). Following Hu and Bentler's (1999) recommendations, we used the following cutoff values of the fit indices as the criteria for an acceptable fit model: (a) CFI equal or larger than 0.95, (b) RMSEA equal to or smaller than 0.08, and (c) SRMR equal or smaller than 0.08. Additionally, the final selected model would be determined by three generally recommended information criteria (Nylund et al., 2007); namely, Akaike information criterion (AIC), Bayesian information criterion (BIC), and the sample size-adjusted Bayesian information criterion (SABIC) for comparing non-nested models. The smaller value of these information criteria indicated a better fit model.

Results

Descriptive statistics

The means and standard deviations of all 10 CLASS dimensions (Pianta et al., 2008) are presented in Table 1 with a breakdown by years of observations (i.e., Y01, Y02, and Y03). As illustrated, the means of all dimensions fell in the middle to high range (2.70–5.78) except for Negative Climate, which was rated in the lower range (1.21–1.87). The patterns of average ratings on the 10 dimensions were consistent across the observation years. Thus, the three cohorts of children were combined as a whole sample for the following analyses ($N = 137$).

The bivariate correlations for the 10 dimensions of CLASS K-3 are presented in Table 2. As illustrated, correlations among the 10 dimensions ranged from −0.29 to 0.90 and were all statistically significant ($p < 0.05$) except for the four correlations related with Negative Climate (Productivity [$r = -0.16$], Instructional Learning Formats [$r = -0.15$], Quality of Feedback [$r = -0.11$], and Language Modeling [$r = -0.16$]).

TABLE 1 Means and standard deviations (in parentheses) for measures of observed classroom interactions.

	Classroom observations		
Year of observation	Y01	Y02	Y03
Number of classrooms observed	42	49	46
Type of observations	Video	Video	Video
Emotional support			
Positive climate	5.35 (0.81)	5.46 (1.59)	5.78 (0.59)
Negative climate	1.87 (1.10)	1.21 (0.42)	1.28 (0.95)
Teacher sensitivity	4.55 (0.89)	5.01 (1.53)	4.92 (1.25)
Regard for student perspectives	3.63 (1.25)	4.65 (1.64)	4.47 (1.67)
Classroom organization			
Behavior management	4.80 (1.07)	4.91 (1.68)	5.21 (1.16)
Productivity	5.27 (0.82)	5.06 (1.46)	5.45 (0.95)
Instructional learning formats	4.54 (1.02)	4.76 (1.41)	5.16 (0.98)
Instructional support			
Concept development	2.70 (1.50)	3.87 (1.73)	3.32 (1.51)
Quality of feedback	3.05 (1.50)	3.66 (1.88)	3.31 (1.51)
Language modeling	3.69 (1.25)	3.91 (1.76)	3.42 (1.38)

Factor structure of the Pre-K classroom dimensions

Results of the three-factor, two-factor, and one-factor CFA models are presented in Table 3, along with three other competing models [i.e., Sandilos and DiPerna's (2014) revised CLASS K-3 model, Pakarinen et al.'s (2010) revised CLASS K-3 model, and Hamre et al.'s (2014) CLASS K-3 revised bi-factor model]. The overall model chi-square statistic, fit indices (CFI, RMSEA, and SRMR) and information criteria (AIC, BIC, and SABIC) are reported for all six models. As shown in Table 3, compared to both two-factor and one-factor models, the three-factor model provided the best overall fit (χ^2 = 119.3, df = 32, $p < 0.001$; CFI = 0.94; RMSEA = 0.14; SRMR = 0.06) with the smallest information criteria. However, even though it fit relatively better than the other two models, the actual fit statistics led to inconsistent conclusion. That is, while SRMR (0.056) was smaller than the corresponding cutoff value, both CFI (0.940) and RMSEA (0.141) fell outside of the recommended cutoff values, and further modification of the three-factor model would be needed to improve model fit. However, we decided not to make further modifications without adequate substantive explanations.

Instead, we decided to fit three other alternative models that were based on previous CLASS K-3 studies. First, Sandilos, DiPerna, and the Family Life Project Key Investigators (Sandilos and DiPerna, 2014) reported a three-factor model with Behavior Management loading on Emotional Support rather than Classroom Organization. Additionally, several residuals were allowed to be correlated with each other (i.e., Negative Climate with Behavior Management, Behavior Management with Productivity, and Regard for Student Perspectives with Concept Development). The resulting model demonstrated a similar fit ($\chi^2(29) = 112.2, \text{CFI} = .94, \text{RMSEA} = .15, \text{SRMR} = .05$) as the three-factor model. Second, we fitted another revised three-factor model (Pakarinen et al., 2010), in which the Negative Climate dimension was omitted and two correlations among the residuals (Behavior Management with Productivity, and Quality of Feedback with Concept Development) were added. Although the information criteria values of the revised Pakarinen model were smaller than alternative models, this was a product of using only nine-dimension measures instead of the 10-dimension measures. Other fit statistics (e.g., RMSEA = 0.17) indicated that this model was not a good fitting model. Lastly, we tested a bi-factor model presented in Hamre et al. (2014). Hamre and colleagues revised the original bi-factor model in two steps. In the first step, the domain-specific factors were allowed to be correlated with each other. In the second step, the revised bi-factor model retained only two domain-specific factors (Positive Management and Routines, and Cognitive Facilitation) while three of the relevant dimensions (Teacher Sensitivity, Regard for Student Perspectives, and Instructional Learning Formats) were only loaded on the general factor to achieve adequate fit.

In the current study, we fitted the first revised bi-factor model with correlated domain-specific factors. As a result, the revised bi-factor model fit the data best among all other alternative models, $\chi^2(22) = 38.1, \text{CFI} = .99, \text{RMSEA} = .07, \text{SRMR} = .02$ (model fit results are highlighted in Table 3). Matching the names of the three original domains proposed by Pianta et al. (2008), we labeled the domain-specific factors Emotional Support, Classroom Organization, and Instructional Support. Although the correlation between the Emotional Support and Classroom Organization factor was high at 0.85 (see Figure 1 for model estimates), we decided not to proceed with the second revised bi-factor model[1] as Hamre et al. (2014) did because it seemed less generalizable to fit to diverse samples. Taken all together, the revised CLASS K-3 bifactor model (Figure 1) was appropriate for our data with adequate fit.

Composite reliability

The Cronbach's alpha coefficients based on the current sample was 0.67 for Emotional Support (α = 0.82 when Negative Climate was recoded to match the direction of scores for Emotional Support), 0.93 for Classroom Organization, 0.95 for Instructional Support, and 0.92 for the total scale. These findings were comparable to those reported by Hamre et al. (2007).

Cronbach's alpha is based on the unidimensionality assumption and the assumption that the items are tau-equivalent measures (Rodriguez et al., 2016). In this study, we found a multidimensional (bi-factor) structure and therefore, omega coefficients can be more appropriate to represent composite reliability (Rodriguez et al., 2016). Different omega coefficients were computed based on the final chosen bi-factor model output. The omega total coefficient was estimated at 0.97 which indicates that the general and domain-specific factors

1 A modified bi-factor model where the Emotional Support and Classroom Organization factors were merged produced worse fit, $\chi^2(25) = 80.69$, CFI = 0.96, RMSEA = 0.13, SRMR = 0.04.

TABLE 2 Bivariate correlations for classroom assessment scoring system K–3 dimensions.

Dimensions	1	2	3	4	5	6	7	8	9	10
1. Positive climate		−0.23**	0.83**	0.66**	0.77**	0.76**	0.71**	0.49**	0.47**	0.46**
2. Negative climate		–	−0.29**	−0.22**	−0.17*	−0.16	−0.15	−0.21*	−0.11	−0.16
3. Teacher sensitivity			–	0.84**	0.83**	0.77**	0.74**	0.65**	0.61**	0.61**
4. Regards for student perspective				–	0.71**	0.66**	0.73**	0.75**	0.70**	0.65**
5. Behavior management					–	0.87**	0.79**	0.57**	0.56**	0.55**
6. Productivity						–	0.79**	0.48**	0.49**	0.49**
7. Instructional learning formats							–	0.69**	0.67**	0.65**
8. Concept development								–	0.90**	0.85**
9. Quality of feedback									–	0.87**
10. Language modeling										–

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

TABLE 3 Model fit indices for the measurement models.

Hypothesized models	Fit indices							
	χ^2	*df*	CFI	RMSEA	SRMR	AIC	BIC	SABIC
1. Three-factor model*	119.3	32	0.940	0.141	0.057	3312.1	3408.5	3304.1
2. Two-factor model	165.0	34	0.909	0.168	0.056	3353.8	3444.3	3346.3
3. One-factor model	432.7	35	0.725	0.288	0.092	3619.5	3707.1	3612.2
4. Sandilos revised three-factor model (Sandilos and DiPerna, 2014)	112.2	29	0.942	0.145	0.050	3311.0	3416.1	3302.2
5. Revised Pakarinen model** (Pakarinen et al., 2010)	105.9	22	0.941	0.167	0.054	2957.6	3051.0	2949.8
6. Revised Bifactor model*** (Hamre et al., 2014)	38.1	22	0.989	0.073	0.022	3250.9	3376.5	3240.4

*The original CLASS K-3 model.
**The revised three-factor model in Pakarinen et al. (Pakarinen et al., 2010) excludes negative climate and use only 9 of the 10 dimensions.
***Associations among the domain factors (Motivational Supports, Management & Routines, and Cognitive Facilitation) were freely estimated while the relationships between Responsive Teaching (overall factor) and the three domain factors were kept being fixed at zero.
Fit indices: RMSEA, Root Mean Square Error of Approximation; CFI, Comparative Fit Index; SRMR, Standardized Root Mean Square Residual; AIC, Akaike information criterion; BIC, Bayesian information criterion; SABIC, Sample-size adjusted BIC (SABIC).

together explained most of the variation in the *total composite score*. Omega hierarchical which shows the variation proportion explained by *only* the general factor out of the *total composite score* variation was estimated at 0.77. As for the omega subscales (or "omega specifics"), estimates were 0.87 (0.90 when Negative Climate was reverse coded), 0.94, and 0.97 for respectively, Emotional Support, Classroom Organization, and Instructional Support. The omega subscales show how much variation in the *subscale composite scores* are explained by the general and domain-specific factors.

Discussion

The current study explored and extended on recent factor analytic studies of the CLASS Pre-K with a sample of Latine preschool-aged children taught by Latine teachers in a Southern state along the United States-Mexico border. As in Hamre et al. (2014), Hu et al. (2016), and more recently Ng et al. (2021), our study provided support for a bifactor model in classrooms serving typically developing Spanish-speaking Latine preschool children taught by Latine teachers.

Consistent with Hamre et al. (2014), a generic factor emerged common to all dimensions in the teaching through interactions framework, and thereby the CLASS Pre-K. Aligned with Hamre et al. (2014) and others, we posit that the generic factor represents a dyadic systems-level property characterized by the construct of responsivity, which functions as a proximal indicator of effective teaching. That is, responsive teachers acknowledge children's verbal and nonverbal communications, support children's autonomy, and display warmth and openness to seeing things from the children's perspective. Responsive teachers also optimize children's learning environments and conditions through communicative strategies that are attuned, concrete, and clear in ways that are developmentally appropriate and promote child engagement (Borairi et al., 2021). Finally, responsive teachers use multiple modalities, including emotional responsivity (e.g., maintaining eye contact, smiling), linguistic responsivity (e.g., recasting children's utterances), and

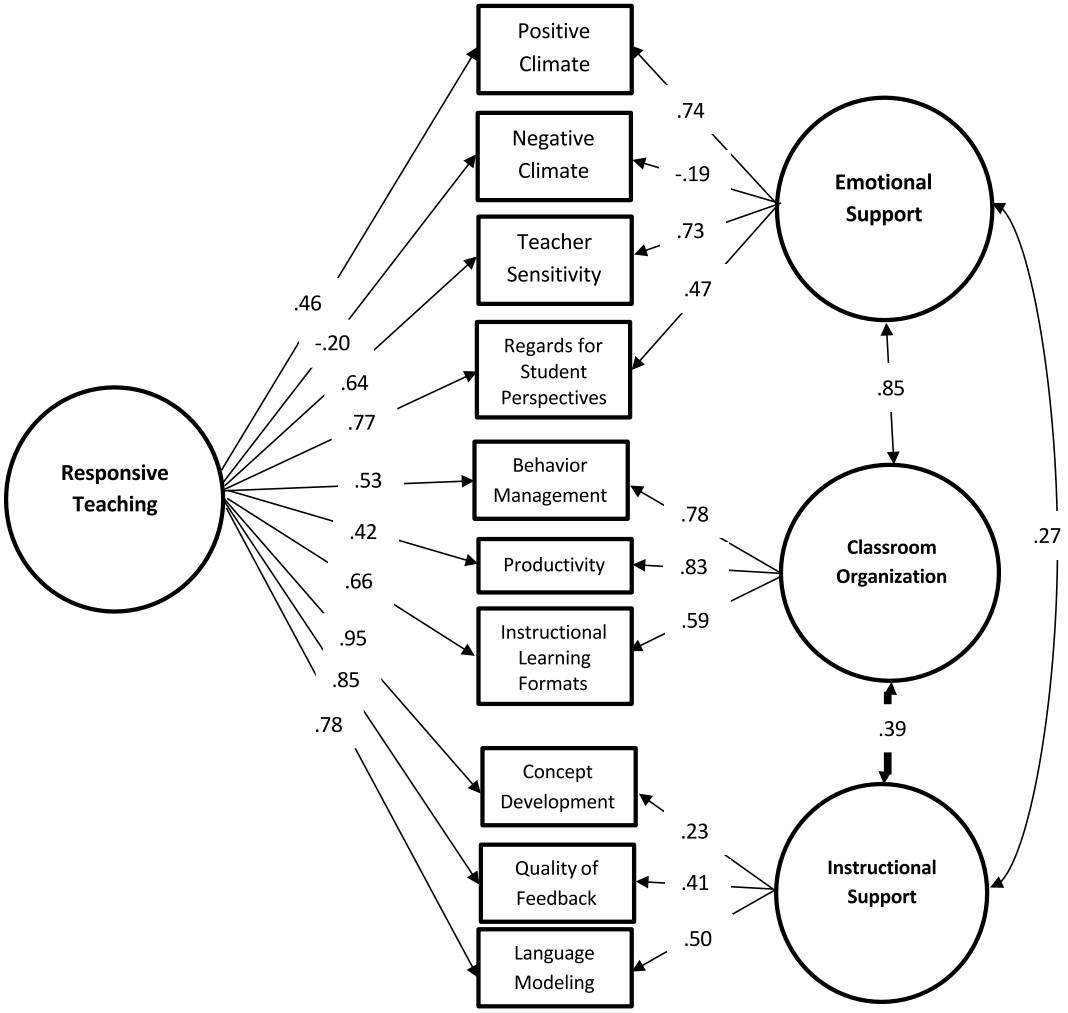

FIGURE 1
Revised classroom assessment scoring system K-3 bi-factor model. All estimates are standardized. The emotional support, classroom organization, and instructional support domain factors were labeled, respectively, as motivational supports, management & routines, and cognitive facilitation in Hamre et al. (2014).

interactive responsivity (e.g., posing open-ended questions), all of which likely play an important role in children's development (Cabell et al., 2011).

The highest loading on Responsive Teaching (RT) was Concept Development whereas Negative Climate had the lowest, yet significant, negative loading. The finding that Concept Development loaded the highest suggests that one of the key underlying attributes of being a responsive teacher is attunement to children's higher-order thinking, as demonstrated by teacher classroom instruction, discussions, and activities with children. Research suggests that very young children can form categorical representations (i.e., concepts) across diverse domains and varying levels of abstraction (Alessandroni and Rodriguez, 2020).

Compared to Hamre et al. (2014), Hu et al. (2016), and Ng et al. (2021), factor loadings for the general factor RT varied across studies. In Hamre et al., the highest factor loading for RT was Teacher Sensitivity, and in both Hu et al. and Ng et al., it was Regard for Student Perspectives (RSP). In the present study, the highest factor loading with RT was, as noted previously, Concept Development (CD). It is unclear why the factor loadings varied across studies. Both Hu et al. (2016) and Ng et al. (2021), which converged on RSP, were conducted in cultural contexts that differ substantively from Western views of early childhood education and child development on which the CLASS is premised. Specifically, Hu et al. was conducted in preschools located in the Guandong province of China whereas the preschools in Ng et al. were located in Singapore. Both countries predominantly adopt teacher-centered approaches to education, with pedagogical ideologies and assumptions deeply rooted in their traditional cultural values. Additionally, the collectivistic nature of Asian societies, which prioritizes conformity and minimizes individualism, significantly shapes educational practices within these contexts (Tan and Rao, 2017). It is reasonable to infer, therefore, as noted by Ng et al. (2021), that behavioral markers as scored in the CLASS are culturally bound to a Westernized perspective.

In contrast to Hamre et al. (2014), Hu et al. (2016), and Ng et al. (2021), whose bifactor models each consisted of a general factor labeled Responsive Teaching and two domain-specific factors labeled Proactive Management and Routines (Positive Management and Routines in Hamre et al., 2014) and Cognitive Facilitation, our study found a generic general factor that we also labeled RT, along with

three, not two, domain-specific factors. The domain-specific factors in our factor model mirrored the original three-factor model theorized by the developers (Hamre et al., 2013). As in the original factor model, we interpret Emotional Support as assessing degree of teacher sensitivity to students' emotional and academic needs in addition to recognition of student autonomy, interests, perspectives, and motivation. Classroom Organization assesses teacher proactive management of student behavior, effective use of time and management of student attention and learning during instruction. Instructional Support evaluates teacher use of approaches to promote higher-order thinking skills, provide feedback to strengthen learning and understanding, and stimulate and facilitate language development (Hamre et al., 2014).

The finding that our results differed from those of other CLASS Pre-K bifactor models is not surprising given that, as noted by Ng et al. (2021), cultural variables likely relate to the nature of teacher-child interactions as behavioral markers are culturally bound and thus affect observations conducted with the CLASS system. Like many of the theories and conceptual ECE frameworks, the "through interactions model" is based on Westernized expectations of teacher-child interactions. However, teacher-child interaction quality is value-laden and culturally loaded such that student behavioral expectations and attendant teacher beliefs may differ substantively across cultural groups and settings (Pastori and Pagani, 2017).

Latine teachers play a key role in the transmission of cultural values, beliefs, traditions, and behavioral norms to Latine students. Although speculative, the transmission likely enhanced the behavioral competencies of children within their own culture of origin, and thus observable characteristics of the classroom environment (Calzada et al., 2010). Our hypothesis is that these values were instantiated in observable daily classroom routines and thus observable by the CLASS Pre-K. The observable behaviors of Latine teachers and their interactions with the Latine preschoolers can reasonably be captured by the Latine values of *bien educado* and *educación*. That is, the constructs reflect the notion of strong moral character, valued social behaviors, book knowledge, and the importance of meeting the social norms and expectations of the community at large (Bridges et al., 2012). It is altogether possible that these values not only guided Latine teacher interactions with their preschool Latine students but also peer interactions, and thus were observable using the CLASS Pre-K.

One interpretation of our findings is that the classroom climate, as measured by the CLASS Pre-K observations in Latine classrooms, captured and reflected the three domains as proposed by the developers. For instance, the Emotional Support domain assesses a teacher's ability to foster and promote emotional functioning within the classroom (Pianta et al., 2008). Relatedly, among Latine cultural values is the central concept of *bien educado,* which includes *cariño,* which is demonstrated in outward expressions of care, emotional warmth, and affectionate behaviors between children, each other, and teachers (Bridges et al., 2012). Latine teachers in our study may have exhibited observable *cariño* by fostering an observable classroom climate characterized by personal connections and communal well-being (Rodriguez et al., 2016).

Classroom Organization examines the effectiveness of routines and procedures in managing children's behavior, time, and attention within the classroom (Pianta et al., 2008). Therefore, it is not unreasonable to theorize that Classroom Organization as observed in the present study was captured by children's demonstration of comportment, cooperation, and obedience in response to the teacher dimensions represented by *bien educado.* That is, it is not surprising that our study replicated the Classroom Organization dimension given that the Latine children in our study classrooms demonstrated observable proper demeanor, respect, obedience, and responsibility as well as warmth in their interactions with each other and in response to teacher efforts at behavior management, structuring classroom routines, and student engagement.

Our study findings also replicated the theorized Instructional Support domain of the CLASS Pre-K. Instructional Support assesses the strategies teachers use to implement the curriculum and promote cognitive and language development (Pianta et al., 2008). It is reasonable to speculate that Instructional Support in the classrooms was instantiated through observable classroom patterns that fused both aspects of the *educación* cultural value and classroom academic and socioemotional expectations. *Educación*, a supraordinate Latine value subsumes *bien educado,* and generally refers to the education of the whole person in terms of book knowledge and learning how to act appropriately within the family and community at large. Thus, it emphasizes the importance of not only academic achievement but also the cultivation of appropriate social behaviors and values within both familial and communal contexts (Bridges et al., 2012).

In summary, this study extends recent factor analytic studies of the CLASS Pre-K by focusing on preschool classrooms with Latine preschoolers taught by Latine teachers. Our study provides support for a bifactor model composed of a general factor labeled Responsive Teaching and three domains aligned with the original CLASS Pre-K factors labeled Emotional Support, Classroom Organization, and Instructional Support. Each of the domains were interpreted in terms of alignment with the Latine cultural values of *bien educado* and *educación* as instantiated in observable classroom routines. Our model adds to the body of research on the cross-cultural applicability of the CLASS both within the United States and internationally. More specifically, the study adds to the accumulating evidence supporting the CLASS's recognition that teacher-child interactions are culturally loaded and bound. In addition, the model fits the "teacher through interactions" theoretical framework as proposed by the developers to construct the CLASS model.

Limitations and future directions

First, the omega subscale coefficient for the Emotional Support scale (0.90) was lower than for the Classroom Organization (0.94), Instructional Support (0.97), and the omega total (0.97) reliability coefficients. However, the composite reliability estimate for Emotional Support was above the expected 0.80 reliability criterion. As previously noted, the Emotional Support scale assesses the degree to which teachers promote a positive classroom climate through their daily interactions with students; that is, how teachers consciously create and promote a sensitive, warm, and trusting environment. It is unclear why this coefficient was substantively lower than the other scales. One possibility is that observable teacher emotional support behaviors across classrooms are far more subtle and culturally loaded than coding for instructional support or classroom organization, thus making ratings somewhat more difficult for the observers to code.

Second, the findings represent point-in-time estimates in school districts populated by large numbers of recent Mexican immigrant families. CLASS Pre-K observations were conducted within a three-day period in October of the project year. While this approach aligns with standard practices for collecting CLASS Pre-K data, it is limited in its ability to capture nuances and fluctuations in classroom interactions and dynamics that may occur over a more extended period.

Third, families and children were not acculturated to U.S. schooling customs. The preschool children of the participating families had been reared with traditional Latine parenting values (e.g., *bien educado*), observable in teacher-child interactions in the classroom. It is altogether possible that our three-domain structure represents unique teacher and child characteristics in the study setting. Although this does not detract from the results, future research is needed to replicate the study in other areas of the country where preschool classrooms populated by Latine populations may be more acculturated or of later generational status (e.g., second, third).

Fourth, our study did not address the predictive validity of the "teaching through interactions" model. Assessing predictive validity is critical given that the domains may not equally predict child outcomes. For instance, Slot et al.'s (2018) research on Danish preschool classrooms indicated limited direct effects of structural or classroom features on children's language development, with only minor associations found with process quality, Emotional Support, and Classroom Organization. Similarly, Gordon and Peng (2020) found that across all CLASS Pre-K domain structures, the correlations with children's academic and socioemotional advancements over a year in Head Start were minimal and largely nonsignificant. Future research focusing on Latine preschool teachers and children should explore the predictive validity of these domain structures on child language, behavioral, emotional, or academic outcomes.

Lastly, our study found a different bifactor structure than previous research. For example, Hamre et al. (2014), Hu et al. (2016), and Ng et al. (2021) noted a bifactor structure with only two domain-specific factors rather than the three domain-specific factors found in the current study. As proposed earlier, this difference may reflect the unique cultural aspects of Latine classrooms, or the findings may be limited to our specific sample. To confirm the findings, therefore, future studies should further investigate the factor structure of CLASS using a similar Latine sample.

Despite these limitations, our study adds to our understanding of the CLASS Pre-K and the influence of cultural contexts on dimensionality. The findings can inform the conceptual basis for the CLASS Pre-K and perhaps further inform theory development leading to nuanced model refinement. Moreover, the finding that our results diverge from other bifactor models and other proposed models gives rise to caution when using the CLASS Pre-K for high-stakes assessment. Specifically, from a practice perspective, it is essential to note that using CLASS Pre-K in classrooms that differ markedly—ethnically or culturally—from the original norming samples may present challenges in interpreting the findings. From a cultural perspective, the original CLASS Pre-K domains may misalign with teacher-child classroom interactions unique to specific cultural groups (e.g., Latine, Chinese), thus risking biased assessments if misinterpreted through the original domains.

Finally, it is noteworthy that our findings, like other studies examining the dimensionality of the CLASS Pre-K, were likely influenced by several unmeasured variables, including variations in observational methodologies, the training and backgrounds of observers, classroom characteristics and composition, as well as the backgrounds and training of teachers. Future research should seek to address these differences to gain a more comprehensive understanding of their influence on the observed outcomes.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by University of Houston Division of Research Human Subjects IRB. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

HK: Formal analysis, Methodology, Writing – review & editing. JG: Conceptualization, Data curation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Funding for this study was provided by Institute of Education Sciences (IES), U.S. Department of Education Grant No. R305A140698 Early Learning Programs and Policies. This material does not necessarily represent the policy of the U.S. Department of Education nor is it necessarily endorsed by the federal government.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Alessandroni, N., and Rodriguez, C. (2020). The development and categorization of conceptual thinking in early childhood: methods and limitations. *Psychol. Res. Rev.* 33, 1–20. doi: 10.1186/s41115-020-00154-9

Bihler, L. M., Agache, A., Kohl, K., Willard, J. A., and Leyendecker, B. (2018). Factor analysis of the classroom assessment scoring system replicates the three-domain structure and reveals no support for the bifactor model in German preschools. *Front. Psychol.* 9, 1–13. doi: 10.3389/fpsyg.2018.01262

Borairi, S., Fearon, P., Madigan, S., and Plamondon, A. (2021). A mediation meta-analysis of the role of maternal responsivity in the association between socioeconomic risk and children's language. *Child Dev.* 92, 2177–2193. doi: 10.1111/cdev.13695

Bridges, M., Cohen, S. R., McGuire, L. W., Yamada, H., Fuller, B., Mireles, L., et al. (2012). Bien Educado: measuring the social behaviors of Mexican American children. *Early Child Res. Q.* 27, 555–567. doi: 10.1016/j.ecresq.2012.01.005

Bronfenbrenner, U., and Morris, P. A. (1988). "The ecology of developmental processes" in Handbook of child psychology: Theoretical models of human development. eds. W. Damon and R. M. Lerner. *5th* ed (New York: John Wiley & Sons, Inc.), 993–1028.

Brown, J. L., Jones, S. M., LaRusso, M. D., and Aber, L. (2010). Improving classroom quality: teacher influences and experimental impacts of the 4Rs program. *J. Educ. Psychol.* 102, 153–167. doi: 10.1037/a0018160

Cabell, S. Q., Justice, L. M., Piasta, S. B., Curenton, S. M., Wiggins, A., Turnbull, K. P., et al. (2011). The impact of teacher responsivity education on preschoolers' language and literacy skills. *Am. J. Speech Lang. Pathol.* 20, 315–330. doi: 10.1044/1058-0360(2011/10-0104)

Calzada, E. J., Fernandez, Y., and Cortes, D. E. (2010). Incorporating the cultural value of respeto into a framework of Latine parenting. *Cult. Divers. Ethnic Minor. Psychol.* 16, 77–86. doi: 10.1037/a0016071

Gordon, R. A., and Peng, F. (2020). Evidence regarding the domains of the CLASS Pre-K in head start classrooms. *Early Child Res. Q.* 53, 23–39. doi: 10.1016/j.ecresq.2020.01.008

Hamre, B. K., Hatfield, B. E., Pianta, R. C., and Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: associations with preschool children's development. *Child Dev.* 85, 1257–1274. doi: 10.1111/cdev.12184

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., et al. (2013). Teaching through interactions: testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *Elem. Sch. J.* 113, 461–487. doi: 10.1086/669616

Hamre, B. K., Pianta, R. C., Mashburn, A. J., and Downer, J. T. (2007). Building a science of classrooms: Application of the CLASS framework in over 4,000 U.S. early childhood and elementary classrooms. New York: Foundation for Child Development.

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Hu, B. Y., Fan, X., Gu, C., and Yang, N. (2016). Applicability of the classroom assessment scoring system in Chinese preschools based on psychometric evidence. *Early Educ. Dev.* 27, 714–734. doi: 10.1080/10409289.2016.1113069

Iruka, I. U., Sheridan, S., Koziol, N., Schumacher, R., Kerby, H., Prokasky, A., et al. (2022). Examining malleable factors that explain the end-of-kindergarten racial/ethnic gaps. *Elem. Sch. J.* 122, 379–410. doi: 10.1086/718072

Karlsen, L., Ryland, V., Buøen, E. S., Lowe Vandell, D., and Lekhal, R. (2024). The factor structures of the classroom assessment scoring system Pre-K and mature play observation tool in multi-ethnic Norwegian early childhood centers. *J. Early Educ. Res.* 22, 106–121. doi: 10.1177/1476718X231195708

Li, H., Liu, J., and Hunter, C. (2020). A meta-analysis of the factor structure of the classroom assessment scoring system (CLASS). *J. Exp. Educ.* 88, 265–287. doi: 10.1080/00220973.2018.1551184

McDoniel, M. E., Townley-Flores, C., Sulik, M. J., and Obradović, J. (2022). Widely used measures of classroom quality are largely unrelated to preschool skill development. *Early Child Res. Q.* 59, 243–253. doi: 10.1016/j.ecresq.2021.12.005

McWayne, C. M., Melzi, G., Limlingan, M. C., and Schick, A. (2016). Ecocultural patterns of family engagement among low-income Latino families of preschool children. *Dev. Psychol.* 52, 1088–1102. doi: 10.1037/a0040343

Muthén, L. K., and Muthén, B. O. (1998–2017). Mplus user's guide. *8th* Edn. Los Angeles, CA: Muthén & Muthén.

Ng, E. L., Bull, R., Bautista, A., and Poon, K. (2021). A bifactor model of the classroom assessment scoring system in preschool and early intervention classrooms in Singapore. *Int. J. Early Child.* 53, 197–218. doi: 10.1007/s13158-021-00292-w

Nylund, K. L., Asparouhov, T., and Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct. Equ. Model.* 14, 535–569. doi: 10.1080/10705510701575396

Pakarinen, E., Lerkkanen, M., Poikkeus, A., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., et al. (2010). A validation of the classroom assessment scoring system in Finnish kindergartens. *Early Educ. Dev.* 21, 95–124. doi: 10.1080/10409280902858764

Pastori, G., and Pagani, V. (2017). Is validation always valid? Cross-cultural complexities of standard-based instruments migrating out of their context. *Eur. Early Child. Res. J.* 25, 682–697. doi: 10.1080/1350293X.2017.1356545

Pianta, R. C., La Paro, K. M., and Hamre, B. K. (2008). The classroom assessment scoring system manual, K-3. Baltimore, MD: Brookes.

Rodriguez, A., Reise, S. P., and Haviland, M. G. (2016). Evaluating bifactor models: calculating and interpreting statistical indices. *Psychol. Methods* 21, 137–150. doi: 10.1037/met0000045

Sandilos, L. E., and DiPerna, J. C. (2014). The measuring quality in kindergarten classrooms: structural analysis of the classroom assessment scoring system (CLASS K–3). *Early Educ. Dev.* 25, 894–914. doi: 10.1080/10409289.2014.883588

Slot, P. L., Bleses, D., Justice, L. M., Markussen-Brown, J., and Højen, A. (2018). Structural and process quality in Danish preschools: direct and indirect associations with children's growth in language and preliteracy. *Early Educ. Dev.* 29, 581–602. doi: 10.1080/10409289.2018.1452494

Tan, T. T., and Rao, N. (2017). How do children learn? Beliefs and practices reported by kindergarten teachers in Singapore. *Early Child Res Assoc.* 11, 81–112. doi: 10.17206/apjrece.2017.11.3.81