Psychol Stud
https://doi.org/10.1007/s12646-019-00488-0

RESEARCH IN PROGRESS

Factorial Validity of the Student–Teacher Relationship Scale—Short Form, Latent Means Comparison of Teacher–Student Relationship Quality and Association with Child Problem and Prosocial Behaviours

Michael Osei Aboagye[1] · Jinliang Qin[1] · Simona Pekárková[2] · Collins Opoku Antwi[3] · Yasin Jababu[4] · Kotor Asare[5] · Emmanuel Affum-Osei[6] · Nila Akinyi[7]

Received: 26 July 2018 / Accepted: 16 April 2019

Abstract The present study verified the hypothesised two-factor structure of the student–teacher relationship scale, short form on Ghanaian sample, using confirmatory factor analysis, multi-group confirmatory factor analysis and structured latent means analysis. On preschool sample (N = 2583; M_{age} = 4.29, SD = 1.34) from 10 regions, the scale's measurement invariance is tested across age, gender, and school types. The confirmatory factor analysis supported the two-factor structure: closeness and conflict. The proposed two-factor model is found to be valid and reliable in the Ghanaian preschool context. Partial strong factorial equivalence across age, gender, and school types was identified in the findings. Significant differences in teacher–child relationship quality were found between boys and girls, and school types. The results provide implications for early childhood education stakeholders in Ghana. Also, it adds to the evidence of cross-cultural applicability of the student–teacher relationship scale—short form.

Keywords Teacher–child relationship · Student–teacher relationship scale (STRS-SF) · Factorial validity · Latent means comparison

Introduction

Studies on teacher–child relationships have evolved to typify complex socio-economic, cultural and psychological expressions of human behaviours broadly. Teacher–child relationships considered from the viewpoint of the ecological transition by Pianta, Hamre, and Stuhlman (2003), paints a picture of "micro-systems that consist of multiple interrelated characteristics and perceptions that both the child and teacher have about contingent interactions". The

✉ Yasin Jababu
yjababu@aggies.ncat.edu

Michael Osei Aboagye
aoseimichaelaboagye@yahoo.com

Jinliang Qin
qinjinlian126@126.com

Simona Pekárková
simona@pekarkova.eu; simona@techsophia.cz

Collins Opoku Antwi
cantwi28@yahoo.com; coantwi@st.usst.edu.cn

Kotor Asare
akotor@uew.edu.gh

Emmanuel Affum-Osei
affumnanaosei@yahoo.com; affumosei@link.cuhk.edu.hk

Nila Akinyi
nakinyi@mmarau.ac.ke

[1] Zhejiang Normal University, Jinhua, China

[2] The University of Hradec Kralove, Hradec Králové, Czech Republic

[3] University of Shanghai for Science and Technology, Shanghai, China

[4] North Carolina A&T State University, 1601 East Market Street, Greensboro, NC 27411, USA

[5] University of Education, Winneba, Ghana

[6] The Chinese University of Hong Kong, Hong Kong, China

[7] Maasai Mara Univiersity, Narok, Kenya

Published online: 11 June 2019

recent development of research on teacher–child relationships has been associated with the interest in assessing the global applicability of the widely used Student–Teacher Relationship Scale—Short Form (STRS-SF) (Drugli & Hjemdal, 2013; Settanni, Longobardi, Sclavo, Fraire, & Prino, 2015; Spilt, Koomen, & Jak, 2012; Whitaker, Dearth-Wesley, & Gooze, 2015). Previous studies have focused on the scale's validity across cultures (Beyazkurk & Kesner, 2005), measurement equivalence (Koomen et al., 2012), and comparison of quality levels of teacher–student relationships across children's gender, age and school context (Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert, 2014).

Interestingly, studies on teacher–student relationship assessment and the validity of the STRS-SF have been conducted in Western countries, assessing the factorial validity in Europe with studies in Greek (Gregoriadis & Nikolaos, 2008), Italian (Fraire, Longobardi, Prino, & Sclavo, 2013a, 2013b), Turkish and American (Beyazkurk & Kesner, 2005), Dutch (Koomen et al., 2012), and the Scandinavian (Norwegian) (Drugli & Hjemdal, 2013) contexts. However, the scale's applicability and factorial validity in African contexts are rarely evaluated, limiting researchers' position on the scale's global applicability.

Due to the limited studies on STRS-SF validity assessment and the lack of research on teacher–student relationship in African contexts, the objectives for undertaking this research are twofold: to assess whether or not (1) the factor structure of the STRS-SF fits the Ghanaian preschool contexts, (2) teacher–student relationship quality differs significantly across children's gender, age, and school type.

Literature

Measuring TCR in Preschools: A Case for Selecting the Modified Version of STRS

The Student–Teacher Relationship Scale (STRS), original and modified versions (Pianta & Steinberg, 1992; Pianta, 2001), are the most validated and widely acceptable tools for measuring teachers' perceptions of social interactions and relationships with their students in preschools (Cadima et al., 2015; Hamre et al., 2008; Zhang & Nurmi, 2012). Previous studies largely offer support for teacher–student relationship quality as a multi-dimensional construct, consisting of closeness, conflict, and dependency components (Fraire, Longobardi, Prino, Sclavo, & Settanni, 2013; Gregoriadis & Nikolaos, 2008; Koomen et al., 2012; R. C. Pianta, 2001; Spilt et al., 2012; Webb & Neuharth-pritchett, 2011).

Generally, previous findings indicate that the original scale has relatively acceptable internal consistency, with Cronbach's alpha coefficients ranging from .91 to .93 (conflict subscale), .85 to .87 (closeness subscale) and .40 to .70 (dependency subscale) (Fraire, Longobardi, Prino, & Sclavo 2013a, 2013b; Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert 2014; Rey et al., 2007; Webb & Neuharth-pritchett, 2011), and construct validity (Buyse et al., 2008; Cadima et al., 2015; Collie et al., 2016; Drugli & Hjemdal, 2013; Fraire, Longobardi, Prino, & Sclavo, 2013a, 2013b; Lessard, Poirier, & Fortin, 2010; Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert, 2014; Settanni et al., 2015). That notwithstanding, noteworthy cautions have been sounded with regards to the use of the original scale in the wake of its inherent limitations such as contextual fitness of the scale at a global level in respect of structural and factorial validity [see (Drugli & Hjemdal, 2013; Gregoriadis & Nikolaos, 2008; Koomen et al., 2012; Webb & Neuharth-pritchett, 2011)], and measurement non-equivalence across age and gender (Koomen et al., 2012; Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert, 2014; Spilt et al., 2012; Webb & Neuharth-pritchett, 2011), and ethnic groups (see Webb & Neuharth-pritchett, 2011). These works sum up the contentions scholars have with the structural validity of the original 28-item STRS.

Earlier findings on the applicability, reliability, and measurement non-invariance of the 28-item version in some geographic contexts stimulated the interest of scholars on the need for further factorial validity and measurement invariance investigations into the STRS globally, culminating in debates on content revision. Consequently, the original 28-item STRS was revised to a 15-item scale (Whitaker et al., 2015), consisting only closeness and conflict constructs of the teacher–child relationship quality. The closeness dimension examines teachers' sensitivity and responsiveness, warmth, affection, and open communication with children (Birch & Ladd, 1997; Pianta & Steinberg, 1992). Conflict, on the other hand, measures the magnitude of negativity in teacher–child relationships. A conflict relationship between a teacher and a child is characterised by discordant interactions, the absence of teacher–child rapport, and low sensitivity on the part of the teacher (Birch & Ladd, 1997; Drugli, 2013).

The dependency subscale, which examines the degree of the child's over-reliant behaviour on the teacher (Birch & Ladd, 1997), was excluded, due to validity and reliability issues (Fraire, Longobardi, Prino, & Sclavo, 2013; Pianta, 2001; Rydell, Bohlin, & Thorell, 2005; Spilt et al., 2012), and its culturally sensitive nature (Beyazkurk & Kesner, 2005; Gregoriadis & Nikolaos, 2008; Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert, 2014; Rydell et al., 2005), given its apparent subjectivity in interpretation (Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert, 2014). For instance, the construct is considered more positive attachment behaviour in close relationship in the

collectivist (Greek and Turkish) cultural contexts (Beyazkurk & Kesner, 2005; Gregoriadis & Nikolaos, 2008), than other countries (Drugli & Hjemdal, 2013).

The STRS-SF has been widely accepted across cultures (Drugli & Hjemdal, 2013; Settanni et al., 2015; Spilt et al., 2012; Whitaker et al., 2015), and received empirical support from teacher–child relationship quality studies, regarding its factorial validity and measurement invariance, mainly in Western collectivist and individualist samples (see Drugli & Hjemdal, 2013; Gregoriadis & Nikolaos, 2008; Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert 2014; Whitaker et al., 2015). In order to strengthen the global applicability of the STRS-SF, its validation in other unexplored cultural contexts has been recommended (Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert 2014).

Comparison of Teacher–Child Relationship Quality Across Students' Gender, Age Groups, and School Types

The potential influence of gender on teacher–child relationship quality and the associated school adjustment, academic problems, or successes can be explored from two vantage points; gender role socialisation and academic risks perspectives (Ewing & Taylor, 2009). On gender role socialisation, differential treatments based on gender which have been found to impact teacher classroom behaviour management systems (Duffy, Warren, & Walsh, 2002; Ewing & Taylor, 2009b; Koch, 2003), reinforce traditional gender stereotypes (Chapman, 2016), and affect teacher–child relationship quality between boys and girls differently (Ewing & Taylor, 2009). Girls are expected to be more strongly impacted by close/positive teacher–student relationship quality given their social/relational gender-typed orientation such as natural propensity for intimacy, affiliation, and emotional connectedness than boys who are expected to assert dominance and independence (Maccoby, 1990).

Academic risk perspective has it that children at greater risk of failure gain or suffer greater benefits or loss in respect of their classroom adaptability (Hamre & Pianta, 2001). Empirical studies show that boys exhibit more problematic academic and behavioural tendencies, which means boys generally exhibit adjustment problems, and are more combatant with their teachers than girls (Birch & Ladd, 1997; Silver, Measelle, Armstrong, & Essex, 2005). Nevertheless, Koomen et al. (2012) suggest that although measurement invariance tests are imperative when comparing boys with girls in teacher–child relationship, the authors lament that extant studies assumed equivalence without empirical tests.

Education, in most part, has always remained a serious government business—ranging from funding, regulation to management, especially at lower levels. Private educational institutions, mainly religious schools like the Catholic private schools have played a huge role in education across Europe and the USA (Coleman et al., 1982a, 1982b; Corten & Dronkers, 2006; Goldhaber, 1996). Studies into differential performance in Europe and the USA among public schools and government-dependent and government-independent private schools have produced consistent results with government-dependent private schools outperforming public schools, and public schools generally outshining government-independent private schools (Corten & Dronkers, 2006; Sassenrath, Croce, & Penaloza, 1984), with few studies as exceptions (even that, sectorial performance differentials along these lines have been observed) (see Witte, 1992). Particularly interesting is the performance outcome in the spheres of cognitive abilities (Coleman et al., 1982a, 1982b). Teacher–child relationship quality, especially closeness component, has been tipped in numerous studies to predict students' academic outcomes and general school likeness (Birch & Ladd, 1997; Hamre & Pianta, 2001; Jerome et al., 2009). With the established academic performance differences in public and private schools, this study intended to test STRS-SF measurement invariance between school types in the Ghanaian preschool context to permit comparisons.

Children's age has been closely researched with respect to behavioural and inter-personal relationship changes and their academic outcomes (Birch & Ladd, 1997, 1998; Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert 2014). Grounded in attachment theories, parent–child relationships exhibit, in most part, a continued stability. However, the nature of children as rational beings is one of adaptability—to changing environment; school, classrooms, teachers among others. Effectively, it is necessary that as children move through preschool or school, their relationship with different teachers along the educational ladder be assessed. As children grow, changing classes almost always come with changing teachers. Expectedly, scholars have witnessed notable changes in teacher–child relationship quality through 1st grade from preschool (see Birch & Ladd, 1998; Pianta & Stuhlman, 2004). Deviating from these findings, are the outcomes discussed by Howes and colleagues, that teacher–child relationship quality does not record significant variation from toddler to preschool years, and assigned children's probable skills for internalising relational styles instead of characteristics of particular teacher, as accounting for the invariance (Howes, 2000; Howes, Hamilton, & Philipsen, 1998). Again, Solheim, Berg-Nielsen, & Wichstrøm (2012) were apprehensive that findings of studies such as Pianta (2001) and Pianta & Stuhlman (2004) concluding on the found

differences of teacher–child relationship quality across different age groups are based on possible erroneous assumption that the core tenets of STRS stay unchanged across ages. Premised on these revelations, it became imperative to test the equivalences of teacher–child relationship quality across nursery and kindergarten age groups in validating the STRS-SF.

Methods

Participants and Procedure

The data used for this study is part of a bigger data on preschool teachers' well-being and teacher–child relationship quality in low- and middle-income countries (Aboagye et al., 2018), funded and supervised by the College of Preschool Teacher Education of Zhejiang Normal University (ZJNU), China. The study was conducted in some selected government and non-government-funded preschools located in the 10 regions of Ghana, using stratified and simple random sampling techniques. In each region, regional and district capitals were selected in addition to some surrounding villages for sampling. The regional and district capitals constitute the sociocultural, as well as techno-economic and administrative centres of the country, with school pupils embodying these diverse backgrounds. Therefore, greater part of regional samples was taken from the regional capitals (n = 1297), followed by the district capitals (n = 713), and then, the villages (n = 573). This in effect made the study sample (N = 2583) a representative of the characteristics of the Ghanaian preschool population.

The schools' classroom compositions are of a typical multi-group classroom character; composed of pupils with different age categories beginning from 2 to 6 years, disaggregated into (1) nursery (children aged from 2 to 4) and kindergarten (children aged above 4 to 6). The final data comprised of boys (n = 1310) and girls (n = 1273) with age ranging from 2 to 6 years (M_{age} = 4.29, SD = 1.34). The pupils were relatively distributed across school types (government-funded: n = 1370 and non-government-funded preschools: n = 1213).

Measures

Student–Teacher Relationship Scale (STRS-SF)

The 15-item student–teacher relationship scale (STRS-SF) (Pianta, 2001) was adopted. It consisted of two subscales: closeness and conflict, rated on five-point scale: "1 = Definitely does not apply to 5 = Definitely applies". The Closeness subscale measures teachers' feelings of affection and open communication with their pupils (e.g. *I share an affectionate, warm relationship with this child*). The conflict subscale, on the other hand, measures teachers' perception of conflict and negativity with the children under their care (e.g. *This Child and I always seem to be struggling with each other*). STRS-SF demonstrated internal consistency reliability with Cronbach's alpha coefficients of .89 and .78 for the closeness and conflict subscales, respectively. Its validity is also reported in previous studies (Beyazkurk & Kesner, 2005; Gregoriadis & Nikolaos, 2008), and supported by the current study (see Table 3).

Strength and Difficulty Questionnaire (SDQ)

Teachers responded to the children and adolescents measure for adjustment and psychopathology (Goodman, 2001) items to rate children's negative emotional and behavioural attributes (difficulties): (1) emotional symptoms (e.g. *worries, clingy*: a = .74), (2) hyperactivity–inattention (e.g. *distractible, restless*: a = .79), (3) peer problems (e.g. *solitary, unpopular:* a = .71), (4) conduct problems (e.g. *fight, lies*: a = .80), and (5) prosocial behaviour (e.g. *helping, caring*: a = .83) for positive attributes (strength). Reliability and validity of the SDQ have been reported (Goodman, 2001; Koomen et al. 2012; Milatz, Glüer, Harwardt-heinecke, Kappler, and Ahnert 2014).

Data Preparation, Descriptive Statistics and Analytic Approach

Missing Data and Descriptive Statistics

Cases with missing data (2.42%) were retained and treated as missing at random (MAR) by using full-information maximum likelihood estimators (FIML) (Allison, 2003). In line with Curran et al. (1996) and Allison (2003), item descriptive statistics (mean and standard deviation), and normality (skewness and kurtosis) were examined. Skewness and kurtosis met their recommended threshold ($\leq \pm 2.0$), indicating normal distribution, a critical assumption underlying the maximum likelihood procedure (Muthen & Kaplan, 1985; Curran, West, & Finch, 1996; Finney & DiStefano, 2006).

Data Analytic Approach

Structural validity was assessed in a two-step strategy: (1) partial confirmatory factor analyses (PCFA) and (2) CFA (Gignac, 2009). The PCFA was performed by computing model fit indices from EFA (Gignac, 2009), while CFA was performed in a structural equation modelling, using Amos (Arbuckle, 2013). Two competing models were

Table 1 Summary of steps in measurement invariance and Model fit cut-off criteria

Models	Constrained parameters	Free parameters	Comparison	Model fit cut-off criteria ($N > 1000$)		
				ΔCFI	ΔRMSEA	ΔSRMR
*M*1. Configural invariance	None	FL + Inter + Res + Var + Cov				
*M*2. Metric invariance	FL	Inter + Res + Var + Cov	*M*1–*M*2	≤ .010	≤ .015	≤ .030
*M*3. Scalar invariance	FL + Inter	Res + Var + Cov	*M*2–*M*3	≤ .010	≤ .015	≤ .010
*M*4. Partial scalar invariance	FL + Inter	Inter(NI) + Res + Var + Cov	*M*2–*M*4	–	–	
*M*5. Residual invariance	FL + Inter + Res	Inter(NI) + Res + Var + Cov + Fmean	*M*4–*M*5	≤ .01	≤ .− 01	≤ .001
*M*6.Structural invariance	FL + Inter + Res + Var + Cov	Inter(NI) + FMean	*M*4–*M*6	≤ .01	≤ .− 01	≤ .001

FL factor loadings, *Inter* item intercepts, *Res* item residual variance, *Var* factor variance, *Cov* factor covariance, *NI* non-invariant items, *Fmean* factor mean

compared: (1) a proposed two-factor model (closeness and conflict) in which the underlying items loaded on the proposed factors and (2) an alternative one-factor model in which all items loaded on a single latent factor. Cross-validation of the models was performed on sample-halves (Sample 2 & Sample 3), as suggested by Arlot & Celisse (2010). Model fit was evaluated, using the Chi-square (χ^2) statistics, standardized root mean square error of residual (SRMR), comparative fit index (CFI), normed fit index NFI, and the root mean square error of approximation (RMSEA). Additionally, the Akaike's information criterion (AIC) was used as criteria to examine model fit comparisons, and the model with lower AIC value had the best fit (Akaike, 1987; Bryk & Raudenbush, 1992). CFI $\geq$ .90, SRMR $\leq$ .08 and RMSEA $\leq$ .08 indicated acceptable model fit (Maccallum et al., 1996; Hu & Bentler, 1999).

Reliability and validity of the scale were examined with convergent and discriminant validity, composite reliability, and Cronbach's alpha for internal consistency (Fornell & Larcker, 1981). Average variance explained (AVE) and their square roots were calculated, examined, and compared with inter-factor correlation. Again, maximum shared variance (MSV) and average shared variance (ASV) were computed and compared to the AVE values obtained. Composite reliability and Cronbach's alpha estimates were compared to minimum threshold (0.70) for social and behavioural science studies (Bagozzi, 1993; Hair et al., 2010).

Multi-group confirmatory factor analyses (MGCFAs) were performed to test measurement invariance (MI) (Meredith, 1993) of the STRS-SF across gender, age, and context (school type). Five invariance types were tested; (1) configural, (2) metric, (3) scalar, (4) residual, and (5) structural invariance (Byrne, Shavelson, & Muthen, 1989; Chen, 2007; W Meredith & Teresi, 2006). Following Cheung & Rensvold (2002) and Chen (2007), ΔCFI, ΔRMSEA and ΔSRMR indices were examined for MI model fit. Considering the sample size ($N > 500$), approximately distributed across groups, for testing metric invariance, a change of $\leq$ .010 in CFI, supplemented by changes of $\leq$ .015 in RMSEA and $\leq$ .030 in SRMR indicated equivalence. For test of scalar invariance and the other restrictive models, ΔCFI $\leq$ .010, ΔRMSEA $\leq$ .015 and ΔSRMR $\leq$ .010 indicated equivalence (Cheung & Rensvold, 2002; Chen, 2007; Sass, 2011) (see Table 1).

Comparisons in the STRS-SF's subscales were made using structured latent means analysis (Aiken, Stein, & Bentler, 1994; Hancock, 1997). One group was set as reference group, treating the others as comparison groups. The latent mean for the reference group was fixed at zero (0), and the latent mean for the comparison group estimated as deviation from the reference group (Dimitrov, 2006). Differences in the latent means of the groups are not influenced by the decision to fix reference group's mean at zero while freely estimating the comparison group's mean (Dimitrov, 2006; Hancock, 1997).

Results

Factorial Validity of the Student–Teacher Relationship Scale (STRS-SF)

PCFA (Gignac, 2009) was performed to examine the two-factor structure of the STRS-SF. Results revealed an

Table 2 Fit of the STRS-SF factor model on Ghanaian samples; result of confirmatory factor analysis

Models	χ2 [df]	χ2/ [df]	NFI	CFI	RMSEA	SRMR	Comparison	Δχ2 [Δdf]	AIC
M_1. 2-Factor proposed	272.97[82]	3.33	.97	.98	.04	.03			348.97
M_2. 1-Factor alternative	1689.04[83]	20.35	.79	.79	.12	.13	M_1–M_2	1416.07[1]***	1763.04
Cross-validation									
Sample 2									
M_{1a}. 2-Factor proposed	161.35[81]	1.99	.96	.98	.04	.03			239.35
M_{2a}. 1-Factor alternative	749.69[82]	9.14	.82	.84	.11	.12	M_{1a}–M_{2a}	588.38[1]***	825.69
Sample 3									
M_{1b}. 2-Factor proposed	166.83[83]	2.04	.96	.98	.04	.033			242.83
M_{2b}. 1-Factor alternative	754.99[84]	9.09	.82	.84	.11	.123	M_{2a}–M_{2b}	588.16[1]***	828.99

DF degree of freedom, *NFI* normed fit index, *CFI* comparative fit index, *RMSE* root mean square error of approximation, *SRMR* standardized root mean square error of residuals, *AIC* Akaike's information criterion

***$p < .001$

Table 3 Descriptive statistics, CFA factor loadings, validity and reliability of the STRS-SF

Items: abbreviated item contents	Overall items statistics				Validity & reliability		
	Mean + SD	Skewness	Kurtosis	CFA/CR	CR/a	AVE	ASV
Closeness					.95/.89	.71 (.84[b])	.26
1. c and I share affectionate, warm relationship.	3.51 ± 1.24	− .193	− .230	.78/–			
2. c will seek comfort from me, if upset.	3.38 ± 1.22	.005	− .611	.82/25.23			
3. c is uncomfortable with physical affection or touch from me[+].	3.43 ± 1.29	.123	− .450	.85/23.23			
4. c values his/her relationship with me.	3.37 ± 1.22	− .027	− .149	.78/22.97			
5. c beams with pride when I praise him/her.	3.72 ± 1.15	− .005	− .537	.68/19.12			
6. c spontaneously share information about himself/herself.	3.46 ± 1.22	.269	− .712	.75/21.01			
7. It is easy to be in tune with what c is feeling.	2.95 ± 1.42	.587	.035	.51/13.55			
8. c openly shares his/her feeling and experience with me	2.95 ± 1.26	.111	− 1.142	.58/15.66			
Conflict					.98/.78	.85 (.92[b])	.25
9. c and I always seem to be struggling with each other.	4.78 ± 0.79	− .103	− .462	.59/11.79			
10. c easily becomes angry with me.	4.70 ± 0.78	.134	− .554	.57/11.29			
11. c remains angry or resistant after being disciplined.	4.87 ± 0.84	.772	.426	.50/10.14			
12. Dealing with c. drains my energy	4.99 ± .91	.496	− .114	.69/–			
13. When c is in a bad mood, I know we're in for a long and difficult day.	5.047 ± .91	.369	− .611	.56/11.38			
14. c's feelings towards me can be unpredictable or change suddenly	5.26 ± 0.84	.349	− .578	.53/10.61			
15. c is sneaky or manipulative with me	5.33 ± 1.07	.236	− .416	.59/10.80			

c—child; CFA standardized estimates and critical ratios (CR) were recorded from the proposed two-factor model in the overall sample

+= reversed coded item

[a] = Cronbach's alpha for internal consistency

[b] = square roots of average variance explained for discriminant validity

acceptable fit for the preliminary two-factor structure (RMR = .04, CFI = .97, NFI = 94, RMSEA = 03) model.

Based on the PCFA results, we proceeded to confirm the two-factor structure by testing two competing models. The proposed two-factor model (M_1: Table 2) fitted the data compared to the alternative one-factor model (M_2: Table 2). The two-factor model had a lower AIC than the one-factor model, indicating a superior fit. Change in the proposed two-factor model and the alternative one-factor model was significant. These results were cross-validated on the sample-halves 2&3. The proposed two-factor model (M_{1a} and M_{1b}; Table 2) fitted the data, while the alternative one-factor model (M_{1b} and M_{2b}; Table 2) did not. Taken together, change in model fit of the proposed two-factor model across

the two samples was not significant, $\Delta\chi^2_{sample2-sample3} = 5.481$, $\Delta df = 2$, p value > 0.5. Change in model fit of the alternative one-factor model across the two samples was also not significant $\Delta\chi^2_{sample2-sample3} = 5.296$, $\Delta df = 2$, p value > 0.5. Correlation pattern of the manifest scale scores was consistent with that reported in previous studies: closeness was negatively associated with conflict (r =−.23, $p < .001$) (see Table 4).

CFA loadings, Validity and Reliability of the STRS-SF and Association with Problem and Prosocial Behaviours

Table 3 shows the CFA loadings, validity and reliability, and inter-factor correlations between the subscales are presented in Table 4. Results revealed constructs' validity, and items were found to be internally consistent. For constructs' discriminant and convergent validity, AVEs exceeded .50 and the square root of AVEs were greater than the correlation coefficient between factors, supplemented by MSV and ASV < AVE values. Composite reliability (CR) estimates and Cronbach's alpha ratios were greater than .70.

Further reliability evidence based on the scale's relations with other variables was achieved. The Pearson correlation patterns and coefficients for the teacher report of child problem and prosocial behaviour on the SDQ and the manifest STRS-SF scales' scores were in the expected directions, and statistically significant. Closeness correlated negatively with all problem behaviours and positively with prosocial behaviour. On the other hand, conflict was positively associated with problem behaviours and negatively with prosocial behaviour (see Table 4).

Measurement Invariance (MI) and Structured Means Analyses of the STRS-SF

To indicate whether the mean differences of groups can be interpreted meaningfully, series of MI tests on the entire sample were conducted and cross-validated on the sample-halves. Sequentially, constraints were imposed on series of hierarchically nested models to test weak, strong and strict factorial invariance.

Table 4 Associations between the teacher–child relationship quality and child problem and prosocial behaviour

Factors	Closeness	Conflict
1. Closeness		
2. Conflict	− .23***	
3. Emotional symptoms	− .22***	.40***
4. Hyperactivity/inattention	− .36***	.53***
5. Peer problems	− .20***	.39***
6. Conduct problems	− .28***	.55***
7. Prosocial behaviour	.47***	− .46***

***$p < .001$

Invariance Across Gender

Results showed good fit for the configural invariance model test (M_1; Table 5). We then imposed constraints on M_1 to test metric invariance (M_2; Table 5). The results revealed good model fit (ΔCFI = .001). Cross-validation on sample 2 (M_{2a}; Table 5) and sample 3 (M_{2b}; Table 5) verified the initial results with $\Delta CFI_{M2a} = .001$ and $\Delta CFI_{M2b} = .002$, respectively.

The scalar invariance model (M_3; Table 5) showed poor fit (ΔCFI = .086). Cross-validation on samples 2&3 (M_{3a} and M_{3b}; Table 5) justified the initial results with $\Delta CFI_{M3a\&M3b} = .086$. Results showed that the additional constraints imposed on the metric invariance model reduced its fit significantly. Item-level analysis (Sass, 2011) revealed that one item of the closeness subscale (STRS1: *I share an affectionate, warm relationship with this child*) and one of the conflict subscale (STRS11: *dealing with this child drains my energy*) lacked invariance when item intercepts were constrained across groups. These findings were further cross-verified on the sample-halves. Intercepts were freely estimated for the two items, and partial invariance was tested. The result showed acceptable fit (ΔCFI = .003). Cross-validation on samples 2&3 (M_{4a} and M_{4b}; Table 5) verified the initial results, ($\Delta CFI_{M4a} = .001$ and $\Delta CFI_{M4b} = .001$). Residual invariance (M_5; Table 5) and structural invariance (M_6; Table 5) were finally tested by imposing additional constraints. Both M_5 and M_6 were supported because the additional constraints did not alter their model fits significantly (M_5: ΔCFI = .001; M_6: ΔCFI = .001). Cross-validation on the split samples verified the initial results ($\Delta CFI_{M5a} = .003$; $\Delta CFI_{M6a} = .003$, and $\Delta CFI_{M5b} = 002$; $\Delta CFI_{M6b} = .002$).

Invariance Across Age Groups

Results showed good model fit for configural invariance model (M_1; Table 6). After imposing constraints on factor loadings, AIC decreased and ΔCFI = .000, indicating strong metric invariance. Cross-validation on the split samples (M_{2a}. and M_{2b}; Table 6) verified the initial results ($\Delta CFI_{M2a} = .001$, and $\Delta CFI_{M2b} = .002$). The model for scalar invariance (M_3; Table 6) was not supported (ΔCFI = .088), sensing non-invariant items in the model. Item-level analysis showed that the same items (STRS1 and STRS11) lacked invariance across age groups. Cross-validation on samples 2&3 (M_{3a} and M_{3b}; Table 6) justified the initial results ($\Delta CFI_{M3a} = .090$; $\Delta CFI_{M3b} = .082$).

Table 5 Test of invariance of the proposed two-factor STRS-SF across gender; result of multi-group confirmatory factor analysis (MGCFA)

Models	χ^2 [df]	χ^2 [df]	TLI	CFI	RMSEA	SRMR	Comparison	ΔCFI	ΔRMSEA	ΔSRMR	AIC
M_1. Configural invariance	324.34[164]	1.978	.967	.981	.028	.0329					536.34
M_2. Metric invariance	331.48[174]	1.883	.970	.98	.026	.0330	M_1–M_2	.001	.002	.0001	519.48
M_3. Scalar invariance	1038.52[191]	5.437	.879	.894	.059	.0547	M_2–M_3	.086	.033	.0217	1196.52
M_4. Partial invariance[a]	327.25[188]	1.741	.980	.983	.024	.029	M_2–M_4	.003	.002	.0040	497.86
M_5. Residual invariance	349.99[206]	1.699	.976	.982	.023	.0333	M_4–M_5	.001	.001	.0043	477.99
M_6. Structural invariance	351.83[206]	1.708	.976	.982	.023	.0343	M_4–M_6	.001	.001	.0053	479.83
Cross-validating MI on samples 2&3											
Sample 2											
M_{1a}. Configural invariance	246.43[166]	1.593	.968	.974	.030	0.032					472.43
M_{2a}. Metric invariance	275.68[178]	1.549	.970	.975	.029	0.033	M_{1a}–M_{2a}	.001	.001	.0010	459.68
M_{3a}. Scalar invariance	619.54[193]	3.210	.879	.889	.059	0.054	M_{2a}–M_{3a}	.086	.030	.0210	773.54
M_{4a}. Partial invariance[a]	282.12[191]	1.477	.968	.974	.027	0.044	M_{2a}–M_{4a}	.001	.011	.0110	440.12
M_{5a}. Residual invariance	296.89[208]	1.427	.977	.977	.026	0.033	M_{4a}–M_{5a}	.003	.011	.0110	420.89
M_{6a}. Structural invariance	297.53[208]	1.430	.976	.977	.026	0.034	M_{4a}–M_{6a}	.003	.001	.0100	421.53
Sample 3											
M_{1b}. Configural invariance	294.64[166]	1.775	.960	.969	.035	0.046					502.64
M_{2b}. Metric invariance	301.81[178]	1.696	.964	.971	.033	0.041	M_{1b}–M_{2b}	.002	.002	.005	485.81
M_{3b}. Scalar invariance	663.853[193]	3.44	.875	.885	.062	0.060	M_{2b}-M_{3b}	.086	.029	.019	817.85
M_{4b}. Partial invariance[a]	282[191]	1.652	.967	.970	.032	0.047	M_{2b}–M_{4b}	.001	.001	.006	473.53
M_{5b}. Residual invariance	321.581[208	1.546	.972	.972	.029	0.046	M_{4b}–M_{5b}	.002	.003	.001	445.58
M_{6b}. Structural invariance	323.227[208]	1.554	.972	.972	.029	0.048	M_{4b}–M_{6b}	.002	.001	.001	447.23

DF degree of freedom, *TLI* Tucker–Lewise Index, *CFI* comparative fit index, *RMSEA* root mean square error of approximation, *SRMR* standardized root mean square error of residuals, *AIC* Akaike's information criterion

[a]Intercepts of the items "STRS1 and STRS11" were freely estimated across gender

The partial scalar invariance (M_4; Table 6) fitted the data well (ΔCFI = .000). Cross-validation on the split samples (M_{4a} and M_{4b}; Table 6) verified the initial results (ΔCFI_{M4a} = .001 and ΔCFI_{M4b} = .001). Residual invariance (M_5; Table 6) and structural invariance (M_6; Table 6) were then tested and found to be fit. Cross-validation on the sample-halves authenticated the initial results (ΔCFI_{M5a} = .002; ΔCFI_{M6a} = .002, and ΔCFI_{M5b} = .001; ΔCFI_{M6b} = .001).

Invariance Across Context

Configural invariance model (M_1; Table 7) fitted the data well across the two contexts: government-funded and non-government-funded preschools. Imposing constraints on factor loadings reduced AIC and ΔCFI = .001, indicating strong metric invariance. Cross-validation on the split samples (M_{1a} and M_{1b}; Table 7) verified the initial results (ΔCFI_{M1a} = .000, and ΔCFI_{M1b} = .002), respectively. After constraining the factor intercepts, AIC increased and ΔCFI = .090, revealing lack of scalar invariance (see M_3; Table 7) for items "STRS1 and STRS11" across contexts. Cross-validation on the split samples justified the results (ΔCFI_{M3a} = .090 and ΔCFI_{M3b} = .088). We then tested partial scalar invariance (M_4; Table 7) which fitted the data well (ΔCFI = .000). Cross-validation on samples 2&3 (M_{4a} and M_{4b}; Table 7) verified the initial results (ΔCFI_{M4a} = .001 and ΔCFI_{M4b} = .001), respectively.

We further tested residual invariance (M_5; Table 7) and structural invariance (M_6; Table 7), respectively. M_5 and M_6 were buttressed by the results, because imposing extra constraints did not affect the models' fit significantly. Cross-validation on the sample-halves affirmed the initial results (ΔCFI_{M5a} = .000; ΔCFI_{M6a} = .001, and ΔCFI_{M5b} = .002; ΔCFI_{M6b} = .002,

Structured Means Comparison of the Teacher–Student Relationship Across Age, Gender and Context (School Type)

Based on the MI test results, we performed structured means analysis to test latent mean differences across gender, age and school type. Results indicated that the comparison group (boys: n = 1310) had a lower score than the reference group (girls: n = 1273) on the closeness subscale

Table 6 Test of invariance of the proposed two-factor STRS-SF across age groups; result of multi-group confirmatory factor analysis (MGCFA)

Models	χ^2 [df]	χ^2 [df]	TLI	CFI	RMSEA	SRMR	Comparison	Δ CFI	ΔRMSEA	ΔSRMR	AIC
M_1. Configural invariance	373.37[166]	2.25	.967	.974	.031	.0310					581.37
M_2. Metric invariance	382.41[178]	2.15	.970	.974	.030	.0316	M_1–M_2	.000	.001	.0006	566.40
M_3. Scalar invariance	1099.13[193]	5.70	.876	.886	.060	.0597	M_2–M_3	.088	.030	.0281	1253.13
M_4. Partial invariance[a]	393.41[192]	2.05	.972	.974	.029	.0317	M_2–M_4	.000	.001	.0001	549.41
M_5. Residual invariance	400.99[208]	1.93	.976	.976	.027	.0319	M_4–M_5	.002	.002	.0002	524.99
M_6. Structural invariance	400.13[208]	1.92	.976	.976	.027	.0326	M_4–M_6	.002	.002	.0009	540.00
Cross-validation of MI on samples 2&3											
Sample 2											
M_{1a}. Configural invariance	260.61[166]	1.57	.969	.975	.030	.0422					468.61
M_{2a}. Metric invariance	271.23[178]	1.52	.971	.976	.029	.0433	M_{1a}–M_{2a}	.001	.001	.0011	455.23
M_{3a}. Scalar invariance	627.74[193]	3.25	.876	.886	.059	.0702	M_{2a}–M_{3a}	.090	.030	.0269	781.74
M_{4a}. Partial invariance[a]	280.97[192]	1.46	.977	.975	.027	.0434	M_{2a}–M_{4a}	.001	.002	.0001	436.97
M_{5a}. Residual invariance	295.53[208]	1.42	.977	.977	.026	.0439	M_{4a}–M_{5a}	.002	.001	.0005	419.53
M_{6a}. Structural invariance	294.78[208]	1.42	.977	.977	.025	.0455	M_{4a}–M_{6a}	.002	.002	.0021	418.78
Sample 3											
M_{1b}. Configural invariance	274.69[166]	1.66	.966	.973	.032	.0405					482.69
M_{2b}. Metric invariance	281.54[178]	1.58	.970	.975	.030	.0412	M_{1a}–M_{2a}	.002	.002	.0007	465.54
M_{3b}. Scalar invariance	630.45[193]	3.27	.883	.893	.059	.0952	M_{2a}–M_{3a}	.082	.029	.0540	784.45
M_{4b}. Partial invariance[a]	298.04[192]	1.55	.972	.974	.029	.0417	M_{2a}–M_{4a}	.001	.001	.0005	454.04
M_{5b}. Residual invariance	308.09[208]	1.48	.975	.975	.027	.0421	M_{4a}–M_{5a}	.001	.002	.0004	432.09
M_{6b}. Structural invariance	310.19[208]	1.49	.975	.975	.028	.0424	M_{4a}–M_{6a}	.001	.001	.0007	434.19

DF degree of freedom, *TLI* Tucker–Lewise Index, *CFI* comparative fit index, *RMSE* root mean square error of approximation, *SRMR* standardized root mean square error of residuals, *AIC* Akaike's information criterion

[a]Intercepts of the items "STRS1 and STRS11" were freely estimated across age groups

and a higher score on the conflict subscale. These differences were significant ($z = -14.57$, $p < .001$, and $z = 7.89$, $p < .001$).

Similarly, the results supported latent mean difference test across age groups. The comparison group (above 4 year to 6: $n = 1235$) had lower latent mean score than the reference group (age from 2 to 4 years: $n = 1348$) on the closeness, and higher latent mean score on the conflict subscales. However, these differences were not significant (CR = − .14, for closeness and CR = .83, for conflict subscales, $p > .05$ and .10).

In addition, the latent mean differences across context showed the comparison group (Government-funded/public preschools: $n = 1370$) to have lower score on the closeness subscale compared to the reference group (Private preschools: $n = 1213$). On the other hand, the comparison group had higher score on the conflict subscale than the reference group. These differences were significant; closeness (CR = − 16.49, $p < .001$), and conflict (CR = 7.78, $p < .001$). These latent means differences were further explored in simple graphs (Figs. 1, 2 and 3).

Discussion

The present study validated the two-dimensional structure: closeness and conflict of the STRS-SF, using CFA and MGCFA. We verified and extended the cross-cultural validity and measurement invariance evidence of the STRS-SF across Ghanaian preschool children's gender, age and context (i.e. school types). Our study breaks the grounds for future validation of the STSR-SF's factor structure and measurement invariance in other African countries to add to the scale's global applicability. The study, employing statistically sophisticated methods, highlights insightful revelations of the factorial validity and measurement equivalence of the STRS-SF.

First, in line with previous studies, the two-factor structure: closeness and conflict of the hypothesised two-factor model of the STRS-SF was verified in this study. The first-order factor model produced acceptable fit on Ghanaian samples. However, findings of the study indicated some items with psychometric problems. Compared with other studies, the items contained in the conflict subscale, produced relatively low factor loadings which serve as a signal of the items' inadequate predictive

Table 7 Test of invariance of the proposed two-factor STRS-SF across context (school type); result of multi-group confirmatory factor analysis (MGCFA

Models	χ^2 [df]	χ^2 [df]	TLI	CFI	RMSEA	SRMR	Comparison	ΔCFI	ΔRMSEA	ΔSRMR	AIC
M_1. Configural invariance	340.46[164]	2.08	.972	.978	.029	.0290					552.46
M_2. Metric invariance	346.19[176]	1.97	.975	.979	.027	.0290	M_1–M_2	.001	.002	.0000	534.19
M_3. Scalar invariance	1075.98[191]	5.63	.878	.889	.060	.0600	M_2–M_3	.090	.033	.0310	1233.98
M_4. Partial invariance	357.64[189]	1.89	.979	.979	.026	.0290	M_2–M_4	.000	.001	.0000	519.64
M_5. Residual invariance[a]	367.06[206]	1.78	.979	.980	.025	.0200	M_4–M_5	.001	.001	.0090	495.06
M_6. Structural invariance	369.43[206]	1.79	.979	.980	.025	.0310	M_4–M_6	.001	.001	.0020	497.43
Cross-validating MI on Samples 2&3											
Sample 2											
M_{1a}. Configural invariance	245.83[164]	1.49	.973	.979	.028	.0386					457.83
M_{2a}. Metric invariance	256.04[176]	1.46	.975	.979	.027	.0394	M_{1a}–M_{2a}	.000	.001	.0008	444.04
M_{3a}. Scalar invariance	615.10[191]	3.22	.878	.889	.059	.0716	M_{2a}–M_{3a}	.090	.032	.0322	773.10
M_{40a}. Partial invariance[a]	265.34[189]	1.40	.978	.980	.025	.0394	M_{2a}–M_{4a}	.001	.002	.0000	427.34
M_{5a}. Residual invariance	282.84[206]	1.37	.980	.980	.024	.0401	M_{4a}–M_{5a}	.000	.001	.0007	410.84
M_{6a}. Structural invariance	284.53[206]	1.38	.979	.979	.024	.0422	M_{4a}–M_{6a}	.001	.001	.0028	412.53
sample 3											
M_{1b}. Configural invariance	285.73[164]	1.74	.962	.970	.034	.0441					497.73
M_{2b}. Metric invariance	292.40[176]	1.66	.966	.972	.032	.0449	M_{1b}–M_{2b}	.002	.002	.0008	480.40
M_{3b}. Scalar invariance	671.56[191]	3.52	.872	.884	.063	.0641	M_{2b}–M_{3b}	.088	.031	.0192	829.56
M_{4b}. Partial invariance[a]	307.89[189]	1.629	.968	.971	.031	.0457	M_{2b}–M_{4b}	.001	.001	.0008	469.89
M_{5b}. Residual invariance	315.88[206]	1.53	.973	.973	.029	.0456	M_{4b}–M_{5b}	.002	.002	.0001	443.88
M_{6b}. Structural invariance	317.45[206]	1.54	.972	.973	.029	.0458	M_{4b}–M_{6b}	.002	.002	.0001	445.45

DF degree of freedom, *TLI* Tucker–Lewise Index, *CFI* comparative fit index, *RMSE* root mean square error of approximation, *SRMR* standardized root mean square error of residuals, *AIC* Akaike's information criterion

[a]Intercepts of the items "STRS1 and STRS11" were freely estimated across groups

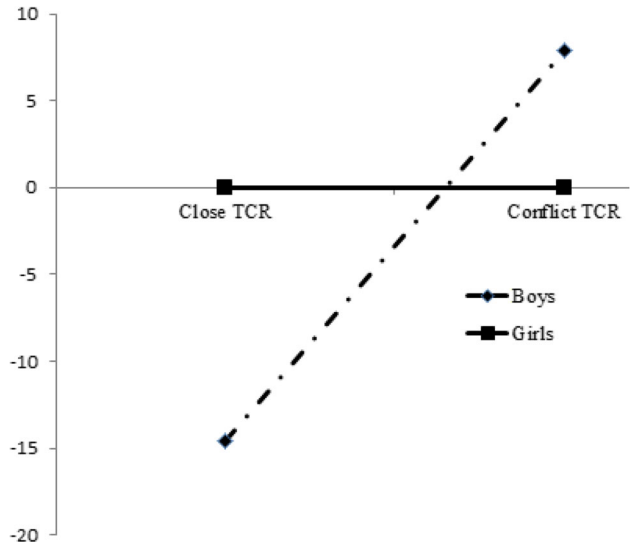

Fig. 1 Patterns of TCR across gender

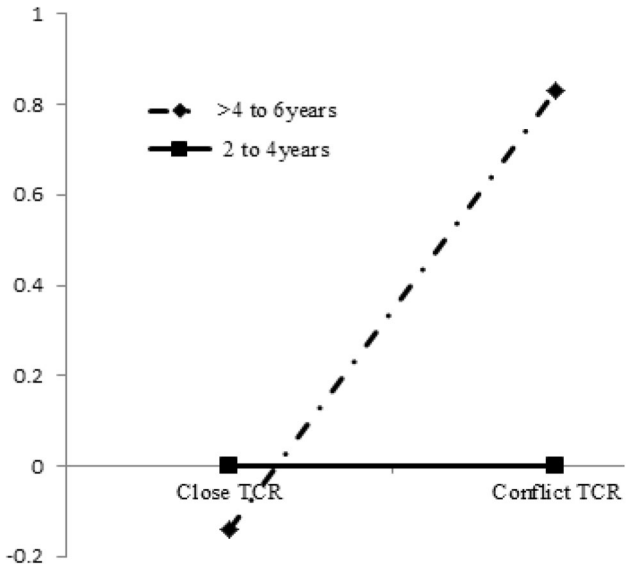

Fig. 2 Patterns of TCR across age

capacity for the latent construct. Nonetheless, the items in the proposed two-factor model showed no model fit discrepancies. Previous studies (e.g. Drugli & Hjemdal, 2013; Whitaker et al., 2015), found these items adequate, in measuring teacher–child relationship conflict; thus, we recommend that future studies explore the concept of

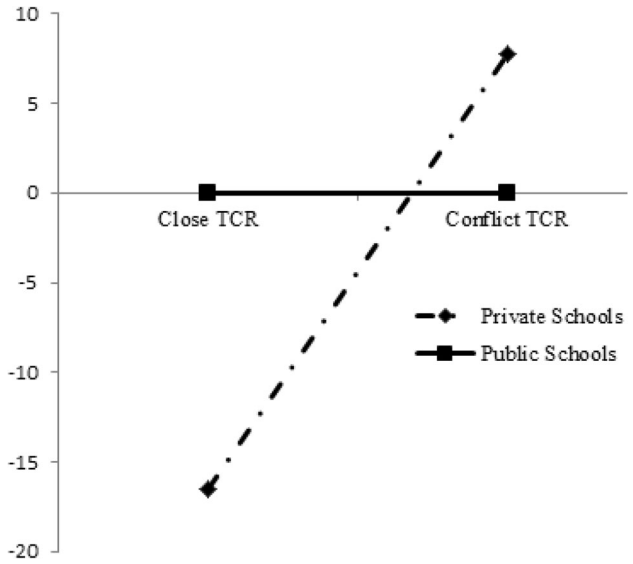

Fig. 3 Patterns of TCR across school type

conflict as conceived specifically in the Ghanaian context, and/or the revalidation of the STRS-SF measures.

Second, the CFA, with the hypothesised two-factor model showed the closeness and conflict factors to have a negative correlation ($r = -23$), confirming the findings of previous studies (Fraire, Longobardi, Prino, & Sclavo, et al., 2013; Webb & Neuharth-pritchett, 2011). The magnitude of the correlation between the factors was considerably lower, compared to the findings of several studies ($r = -.44$ in Gregoriadis & Nikolaos, 2008; $-.62$, in Webb & Neuharth-Pritchett, 2011; $-.35$ in Fraire, Longobardi, Prino, Sclavo, et al., 2013; Settanni et al., 2015), but relatively closer to ($r = -.20$ in Ewing & Taylor, 2009). Our findings also confirmed the discriminant and convergent validity, composite reliability and internal consistency of the STRS-SF, when adopted in Ghanaian context. The result is considerably similar to those of previous studies (e.g. Drugli & Hjemdal, 2013; Hamre & Pianta, 2001; Whitaker et al., 2015). Further support for the STRS-SF's criterion validity was achieved as the manifest scales correlated significantly with problem and prosocial behaviours on the SDQ. Thus, relational conflict appeared to be influenced more by the presence of children's problem behaviours (difficulties) and prosocial behaviours (strengths), confirming previous studies (Koomen et al., 2012; Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert 2014).

The measurement invariance test based on the proposed model in the whole sample, and the sample-halves revealed the items of the STRS-SF to have partial strong factorial invariance across the age and gender of preschool children, and school context (preschool types). Our study indicated non-invariance of items (STRS1: on the closeness subscale *I share an affectionate, warm relationship with this child* and STRS11: on the conflict subscale *dealing with this child drains my energy*) across study groups. For example, teachers' perception of these two items varied among those in nursery and kindergarten. Possible reasons for these variations might be attributable to contextual elements of borderline distinctiveness in semantics, content, technicality and concepts as advised by extant studies (Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert 2014). That notwithstanding, our study results, similar to those of previous studies (Koomen et al., 2012; Milatz, Glüer, Harwardt-heinecke, Kappler, & Ahnert 2014), provide the pre-requisites for multi-group comparison (Meredith & Teresi, 2006), enabling researchers to discern differences in teacher–child relationship quality in respect of the subscales explored (Cheung & Rensvold, 2002; Meredith & Teresi, 2006; Chen, 2007; Sass, 2011).

It is essentially noteworthy that, government-funded and non-government (private and public) preschools' teacher–child relationship quality outcomes were dissimilar, making this a potential information input in education policy formulation and direction. These private–public schools' performance differentials have been noted in prior studies including Coleman et al. (1982a, 1982b). Future enquiries into the factors underscoring these findings may take into account teachers' expectations, patterns of communication and inter-personal socialisation skills, as well as students' socio-economic status (Brady et al., 1992).

Our study found significant differences in teacher–child relationship quality across children's gender. Ghanaian teachers in both public and private preschools perceived a lower sense of relational conflict with girls than with boys. This revelation is in accord with findings of previous studies (e.g. Birch & Ladd, 1997; Howes, 2000; Drugli, 2013), confirming the findings of other studies that girls often have closer relationships with their teachers than boys (Howes 2000). This pattern has been likened to the gender role socialisation perspective (Ewing & Taylor, 2009), which indicates that differential treatments based on gender impact teacher classroom behaviour management systems (Duffy et al. 2002; Koch, 2003; Ewing & Taylor, 2009). These systems reinforce traditional gender stereotypes (Chapman, 2016), which affect teacher–child relationship quality between boys and girls differently (Ewing & Taylor 2009). Thus, girls are more strongly impacted by close/positive teacher–student relationship given their social/relational gender-typed orientation such as natural propensity for intimacy, affiliation and emotional connectedness than boys who are expected to assert dominance and independence (Maccoby, 1990).

Differences in teacher–child relationships across children's age groups were not significant in our study. This result is consistent with those intimated by Howes et al.

(1998). However, several other studies produced results that reflect the developmental perspective (Kopp, 1989; Pianta, 1994; Pianta & Stuhlman, 2004), which states that developmental changes in children typically orchestrate a shift from relevant adults' control to autonomy (Jerome et al., 2009). That notwithstanding, the support for relative stability in teacher–child relationship from toddler to preschool years found by Howes et al. (1998) has been linked to children's internalisation of relational styles instead of teacher characteristics at the preschool stage.

Summarily, the results support the structural validity and factorial equivalence of the STRS-SF in Ghanaian preschools. The evidence of the psychometric properties of the scale corroborates the notion that the STRS-SF is usable, reliable and valid for measuring conflict and closeness constructs of teacher–child relationship quality across age, gender and school types. However, the results of this study should be interpreted within the context of some limitations.

Limitations

The study was conducted mainly within Ghanaian context. Thus, the findings of this study cannot be generalised across other African countries. The factorial invariance test with a specific age range (2–6 years) may limit our findings to the age categories used, which makes it difficult for us to indicate whether the STRS-SF is invariant on children aged above six (6) years. This challenges the generalisability of the results across age categories above 6 years. Regardless of these limitations, this study offers practical implications for ECE management.

Practical Implications for Early Childhood Education

Quality teacher–child relationship creates environments conducive for teaching and learning, and play activities, all of which promote the holistic development of the child. Such undertakings must not be glossed over; they must be made part of the teaching and learning practices, and arrangements in the Ghanaian preschools. These, therefore, call for proper assessment of levels of teacher–child relationship quality using the STRS-SF, in order to shine light on various relationship trajectories necessary or detrimental to children's holistic development.

The need for gender norms in children's assessments must be recognised (Spilt et al., 2012), as differences in latent mean scores for teacher–child relationship quality are evident in the current study. In the context of psychological assessment, we recommend preschool psychologists to use gender-specific norms, for each child.

On policy development, the findings of this study can shape policy discussions, and formulation in creating harmonious educational environment through close study of the differences that spell out the disparate public and private preschools' teacher–child relationship quality. These investigations have become necessary in the light of the findings relating to teacher–child relationship quality differentials in school type.

The findings also add to the previous literature on the theoretical development and cross-cultural assessment of teacher–child relationship quality with the STRS-SF. The results provoke discussions on pressing matters, including reasons why quality teacher–child relationships in Ghanaian preschools matter, in fostering quality early childhood education delivery, with total consideration of gender and age mixes, and school types.

References

Aboagye, M. O., Qin, J., Qayyum, A., Antwi, C. O., Jababu, Y., & Affum-Osei, E. (2018). Teacher burnout in pre-schools: A cross-cultural factorial validity, measurement invariance and latent mean comparison of the Maslach Burnout Inventory, Educators Survey (MBI-ES). *Children and Youth Services Review, 94,* 186–197. https://doi.org/10.1016/j.childyouth.2018.09.041.

Aiken, L. S., Stein, J. A., & Bentler, P. M. (1994). Structural equation analyses of clinical subpopulation differences and comparative treatment outcomes: Characterizing the daily lives of drug addicts. *Journal of Consulting and Clinical Psychology, 62*(3), 488–499.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika 52*(3), 317–332. https://doi.org/10.1007/BF02294359.

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology, 112*(4), 545–557. https://doi.org/10.1037/0021-843X.112.4.545.

Arbuckle, J. L. (2013). IBM SPSS Amos 22 User's Guide IBM Corp, 673.

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys, 4,* 40–79. https://doi.org/10.1214/09-SS054.

Bagozzi, R. P. (1993). Assessing construct validity in personality research: Applications to measures of self-esteem. *Journal of Reasearch in Personality, 27*(6), 49–87.

Beyazkurk, D., & Kesner, J. E. (2005). Teacher-child relationship in Turkish and United States school: A cross-cultural study. *International Education Journal, 6*(5), 554–574.

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology, 35*(1), 61–79. https://doi.org/10.1016/S0022-4405(96)00029-5.

Birch, S. H., & Ladd, G. W. (1998). Children's interpersonal behaviors and the teacher-child relationship. *Developmental Psychology, 34*(5), 934–946. https://doi.org/10.1037/0012-1649.34.5.934.

Brady, B. A., Tucker, C. M., Harris, Y. R., & Tribble, I. (1992). Association of academic achievement with behavior among

black students and white students. *Journal of Educational Research Research, 86*(1), 43–51. https://doi.org/10.1080/00220671.1992.9941826.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models in social and behavioral research: Applications and data analysis methods*. Newbury Park: Sage. Buys.

Buyse, E., Verschueren, K., Doumen, S., Van Damme, J., & Maes, F. (2008). Classroom problem behavior and teacher-child relationships in kindergarten: The moderating role of classroom climate. *Journal of School Psychology, 46*(4), 367–391. https://doi.org/10.1016/j.jsp.2007.06.009.

Byrne, M. B., Shavelson, R. J., & Muthen, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3) 456–466.

Cadima, J., Doumen, S., Verschueren, K., & Buyse, E. (2015). Child engagement in the transition to school: Contributions of self-regulation, teacher–child relationships and classroom climate. *Early Childhood Research Quarterly*. https://doi.org/10.1016/j.ecresq.2015.01.008.

Chapman, R. (2016). A case study of gendered play in preschools: how early childhood educators' perceptions of gender influence children's play. *Early Child Development and Care, 186*(8), 1271–1284. https://doi.org/10.1080/03004430.2015.1089435.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 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.

Coleman, J., Hoffer, T., & Kilgore, S. (1982a). Cognitive outcomes in public and private schools. *Sociology of Education, 55*(2), 65–76. https://doi.org/10.2307/2112288.

Coleman, J. S., Kilgore, S. B., & Hoffer, T. (1982b). *Public and private schools. Society, 19*(2), 4–9. https://doi.org/10.1007/BF02712901.

Collie, R. J., Martin, A. J., Papworth, B., & Ginns, P. (2016). Students' interpersonal relationships, personal best (PB) goals, and academic engagement. *Learning and Individual Differences, 45,* 65–76. https://doi.org/10.1016/j.lindif.2015.12.002.

Corten, R., & Dronkers, J. (2006). School achievement of pupils from the lower strata in public, private government-dependent and private government-independent schools: A cross-national test of the Coleman-Hoffer thesis. *Educational Research and Evaluation, 12*(2), 179–208. https://doi.org/10.1080/13803610600587032.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*(1), 16–29. https://doi.org/10.1037/1082-989X.1.1.16.

Dimitrov, D. M. (2006). Comparing groups on latent variables: A structural equation modeling approach. *Work, 26*(2006), 429–436.

Drugli, M. B. (2013). How are closeness and conflict in student—teacher relationships associated with demographic factors, school functioning and mental health in norwegian schoolchildren aged 6–13? *Scandinavian Journal of Educational Research, 57*(2), 217–225. https://doi.org/10.1080/00313831.2012.656276.

Drugli, M. B., & Hjemdal, O. (2013). Factor structure of the student-teacher relationship scale for norwegian school-age children explored with confirmatory factor analysis. *Scandinavian Journal of Educational Research, 57*(5), 457–466. https://doi.org/10.1080/00313831.2012.656697.

Duffy, J., Warren, K., & Walsh, M. (2002). Classroom interactions: gender of teacher, gender of student, and classroom subject, *45*(November 2001), 579–580.

Ewing, A. R., & Taylor, A. R. (2009). The role of child gender and ethnicity in teacher-child relationship quality and children's behavioral adjustment in preschool. *Early Childhood Research Quarterly, 24*(1), 92–105. https://doi.org/10.1016/j.ecresq.2008.09.002.

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Greenwich: Information Age.

Fornell, C., & Larcker, D. F. (1981). structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 48*(1), 39–50.

Fraire, M., Longobardi, C., Prino, L. E., & Sclavo, E. (2013). La escala " Student-Teacher Relationship Scale" dentro del contexto italiano: Un estudio de validez factorial, *11*(31), 851–882.

Fraire, M., Longobardi, C., Prino, L. E., Sclavo, E., & Settanni, M. (2013b). Examining the student-teacher relationship scale in the italian context: A factorial validity study. *Electronic Journal of Research in Educational Psychology, 11*(3), 851–882. https://doi.org/10.14204/ejrep.31.13105.

Gignac, G. E. (2009). Partial confirmatory factor analysis: Described and illustrated on the NEO-PI-R. *Journal of Personality Assessment, 91*(1), 40–47. https://doi.org/10.1080/00223890802484126.

Goldhaber, D. A. N. D. (1996). Public and private high schools: Is school choice an answer to the productivity problem?*15*(2), 93–109.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337–1345. https://doi.org/10.1097/00004583-200111000-00015.

Gregoriadis, A., & Nikolaos, T. (2008). Applicability of the student—Teacher relationship scale (STRS) in the Greek educational setting. *Journal of Psychoeducational, 26*(2), 108–120.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River: Prentice-Hall Inc.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development, 72*(2), 625–638. https://doi.org/10.1111/1467-8624.00301.

Hamre, B. K., Pianta, R. C., Downer, J. T., & Mashburn, A. J. (2008). Teachers' perceptions of conflict with young students: Looking beyond problem behaviors. *Social Development*. https://doi.org/10.1111/j.1467-9507.2007.00418.x.

Hancock, G. R. (1997). Structural equation modeling methods of hypothesis testing of latent variable means. *Measurement and Evaluation in Counseling and Development, 30*(2) 91–105.

Howes, C. (2000). Social-emotional classroom climate in child care, child-teacher relationships and children's second grade peer relations. *Social Development, 9*(2), 191–204. https://doi.org/10.1111/1467-9507.00119.

Howes, C., Hamilton, C. E., & Philipsen, L. C. (1998). Stability and continuity of child-caregiver and child-peer relationships. *Child Development, 69*(2), 418–426.

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

Jerome, E. M., Hamre, B. K., & Pianta, R. C. (2009). Teacher-Child Relationships from Kindergarten to Sixth Grade: Early childhood predictors of teacher-perceived conflict and closeness. *Social Development*. *18*(4), 915–945. https://doi.org/10.1111/j.1467-9507.2008.00508.x.

Koch, J. (2003). Gender issues in the classroom. In W. M. Reynolds & G. E. Miller (Eds.), *Educational psychology Handbook of psychology*. Hoboken: Wiley.

Koomen, H. M. Y., Verschueren, K., van Schooten, E., Jak, S., & Pianta, R. C. (2012). Validating the student-teacher relationship scale: Testing factor structure and measurement invariance across child gender and age in a Dutch sample. *Journal of School Psychology, 50*(2), 215–234. https://doi.org/10.1016/j.jsp.2011.09.001.

Kopp, C. B. (1989). Regulation of distress and negative emotions: A developmental view. *Developmental Psychology, 25*(3), 343–354. https://doi.org/10.1037/0012-1649.25.3.343.

Lessard, A., Poirier, M., & Fortin, L. (2010). Student-teacher relationship: A protective factor against school dropout. *Procedia Social and Behavioral Sciences, 2*(2010), 1636–1643. https://doi.org/10.1016/j.sbspro.2010.03.250.

Maccallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling of fit involving a particular measure of model. *Psychological Methods, 13*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130.

Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist, 45*(4), 513–520. https://doi.org/10.1037/0003-066X.45.4.513.

Meredith, W. (1993). Measurement Invariance, Factor Analysis and Factorial Invariance. *Psychometrika, 58*(4), 525–543. https://doi.org/10.1007/BF02294825.

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care, 44*(11), S69–S77. https://doi.org/10.1097/01.mlr.0000245438.73837.89.

Milatz, A., Glüer, M., Harwardt-heinecke, E., Kappler, G., & Ahnert, L. (2014). The Student – Teacher Relationship Scale revisited: Testing factorial structure, measurement invariance and validity criteria inGerman-speaking samples. *Early Childhood Research Quarterly, 29*(3), 357–368. https://doi.org/10.1016/j.ecresq.2014.04.003.

Muthen, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor-analysis of non-normal likert variables. *British Journal of Mathematical and Statistical Psychology, 38*(NOV), 171–189. https://doi.org/10.1111/j.2044-8317.1985.tb00832.x.

Pianta, R. C. (1994). Patterns of relationships between children and kindergarten teachers. *Journal of School Psychology, 32*(1), 15–31. https://doi.org/10.1016/0022-4405(94)90026-4.

Pianta, R. C. (2001). Student-teacher relationship scale (STRS): Professional manual.

Pianta, R.C., Hamre, B., & Stuhlman, M. (2003). Relationships between teachers and children. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology* (Vol. 7, pp. 625–653). Hoboken, New Jersey: Wiley.

Pianta, R. C., & Steinberg, M. (1992). Teacher-child relationships and the process of adjusting to school. *New Directions for Child and Adolescent Development, 57,* 61–80. https://doi.org/10.1002/cd.23219925706.

Pianta, R. C., & Stuhlman, M. W. (2004a). Teacher-child relationships and children's success in the first years of school. *School Psychology Review, 33*(3), 444–458. https://doi.org/10.1037/10314-000.

Pianta, R., & Stuhlman, M. W. (2004b). Conceptualizing risk in relational terms: Associations among the quality of child-adult relationships prior to school entry and children's developmental outcomes in first grade. *Educational and School Psychology, 21,* 32–45.

Rey, R. B., Smith, A. L., Yoon, J., Somers, C., & Barnett, D. (2007). Relationships between teachers and Urban African American children the role of informant. *School Psychology International, 28*(3), 346–364. https://doi.org/10.1177/0143034307078545.

Rydell, A. M., Bohlin, G., & Thorell, L. B. (2005). Representations of attachment to parents and shyness as predictors of children's relationships with teachers and peer competence in preschool. *Attachment and Human Development, 7*(2), 187–204. https://doi.org/10.1080/14616730500134282.

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 29*(4), 347–363. https://doi.org/10.1177/0734282911406661.

Sassenrath, J., Croce, M., & Penaloza, M. (1984). Private and public school students: Longitudinal achievement differences? *American Educational Research Journal, 21*(3), 557–563. https://doi.org/10.3102/00028312021003557.

Settanni, M., Longobardi, C., Sclavo, E., Fraire, M., & Prino, L. E. (2015). Development and psychometric analysis of the student—teacher relationship scale—short form. *Frontiers in Psychology, 6*(June), 1–7. https://doi.org/10.3389/fpsyg.2015.00898.

Silver, R. B., Measelle, J. R., Armstrong, J. M., & Essex, M. J. (2005). Trajectories of classroom externalizing behavior: Contributions of child characteristics, family characteristics, and the teacher-child relationship during the school transition. *Journal of School Psychology, 43*(1), 39–60. https://doi.org/10.1016/j.jsp.2004.11.003.

Solheim, E., Berg-Nielsen, T. S., & Wichstrøm, L. (2012). The three dimensions of the student-teacher relationship scale: CFA validation in a preschool sample. *Journal of Psychoeducational Assessment, 30*(3), 250–263. https://doi.org/10.1177/0734282911423356.

Spilt, J. L., Koomen, H. M. Y., & Jak, S. (2012). Are boys better off with male and girls with female teachers? A multilevel investigation of measurement invariance and gender match in teacher-student relationship quality. *Journal of School Psychology, 50*(3), 363–378. https://doi.org/10.1016/j.jsp.2011.12.002.

Webb, M. L., & Neuharth-pritchett, S. (2011). Examining factorial validity and measurement invariance of the Student—Teacher relationship scale. *Early Childhood Research Quarterly, 26,* 205–215. https://doi.org/10.1016/j.ecresq.2010.09.004.

Whitaker, R. C., Dearth-Wesley, T., & Gooze, R. A. (2015). Workplace stress and the quality of teacher-children relationships in Head Start. *Early Childhood Research Quarterly, 30,* 57–69. https://doi.org/10.1016/j.ecresq.2014.08.008.

Witte, F. J. (1992). Private school versus public school achievement: Are there findings that should affect the educational choice debate?*11*(4), 371–394.

Zhang, X., & Nurmi, J. (2012). Teacher – child relationships and social competence: A two-year longitudinal study of Chinese preschoolers. *Journal of Applied Developmental Psychology, 33*(3), 125–135. https://doi.org/10.1016/j.appdev.2012.03.001.