Psychometric Properties of the Student-Teacher Relationship Scale-Short Form in a Norwegian Early Childhood Education and Care Context

Kenneth Stensen[1] ⬤, Stian Lydersen[1], Ingunn Ranøyen[1],
Christian A. Klöckner[2], Elisabet S. Buøen[3,4], Ratib Lekhal[2,5], and
May Britt Drugli[1,6]

Abstract

The Student-Teacher Relationship Scale-Short Form (STRS-SF) is one of the most frequently used instruments globally to measure professional caregivers' perceptions of the relationship quality with a specific child. However, its psychometric properties for children younger than 3 years of age enrolled in early childhood education and care (ECEC) centers are largely unknown. Thus, this study aimed to investigate and evaluate the factorial validity of the STRS-SF and measurement invariance across children's gender and age by combining two large Norwegian community samples (N = 2900), covering the full age range of children enrolled in ECEC (1–6 years olds). Our findings indicate promising psychometric properties for the STRS-SF; thus, its applicability is supported for both younger and older children indiscriminate of their gender. However, some caution is advised when comparing latent means between older and younger ECEC children because professional caregivers interpret the STRS-SF differently based on children's age.

[1]Regional Centre for Child and Youth Mental Health and Child Welfare, Faculty of Medicine and Health Science, Department of Mental Health, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[2]Department of Psychology, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[3]Department of Communication and Culture, BI Norwegian Business School, Oslo, Norway
[4]Regional Centre for Child and Adolescent Mental Health Eastern and Southern Norway, Oslo, Norway
[5]Department of Education, University of Oslo, Oslo, Norway
[6]Centre of the Study of Educational Practice, Inland Norway University of Applied Sciences, Hamar, Norway

Corresponding Author:
K. Stensen, Regional Centre for Child and Youth Mental Health and Child Welfare, Faculty of Medicine and Health Science, Department of Mental Health, Norwegian University of Science and Technology (NTNU), P.O. Box 8905, MTFS, Trondheim N-7491 Trondheim.
Email: kenneth.stensen@ntnu.no

Keywords

student-teacher relationship scale-short form, psychometric properties, early childhood education and care, professional caregivers

Introduction

Healthy development depends on the quality of young children's relationships with significant people in their life. The relationship quality between children and adults established in children's early years lays the foundation for future developmental outcomes, such as mental health and academic achievement (Center on the Developing Child, 2010, 2009; Miller-Lewis et al., 2014). When young children transit from home to early childhood education and care (ECEC) centers, professional caregivers may act as ad hoc attachment figures, performing many similar caregiving functions as parents (Verschueren & Koomen, 2021; Zhang & Sun, 2011). These early bonds may be especially important for the development of very young and vulnerable children, as their capacity for self-regulation is relatively limited and requires adult support (Verschueren & Koomen, 2012). Additionally, teacher-child relationship quality appears to have a moderating effect that may compensate for early negative experiences and reciprocally influences a range of developmental outcomes (e.g., Griggs et al., 2009; Graziano et al., 2016; Sabol & Pianta, 2012; Schmitt et al., 2012; Skalická et al., 2015; Zhang & Sun, 2011).

The student-teacher relationship scale (STRS) (Pianta, 2001) is a widely used instrument to measure professional caregivers and elementary school teachers' perceptions of the relationship quality with a specific child from the age of four and above. However, as more than a third of children in Organisation for Economic Co-operation and Development (OECD) countries attend ECEC centers (OECD, 2021), the instrument is frequently applied for children younger than 4 years old as well. The STRS consists of 28 items that measure three factors: *closeness*, *conflict*, and *dependency*. *Closeness* reflects the professional caregiver's perception of openness, warmth, and security in the professional caregiver-child relationship. *Conflict* refers to the negative perception of the relationship with discordance and unpredictability, while *dependency* refers to the perception of children as developmentally overreliant and possessive (Pianta, 2001).

Several researchers have highlighted reliability and validity issues with the *dependency* factor, suggesting that it may be more susceptible to cultural context and subjective interpretation (Beyazkurk & Kesner, 2005; Doumen et al., 2009; Drugli & Hjemdal, 2013; Fraire et al., 2013; Solheim et al., 2012; Tsigilis & Gregoriadis, 2008; Tsigilis et al., 2018). Consequently, the STRS-short form (STRS-SF) (Pianta, 2001), which contains 15 items measuring the *closeness* and *conflict* factors *(dependency* excluded from the original version), has been frequently applied by researchers investigating professional caregiver-child relationship quality. This two-factor model has shown satisfactory psychometric properties in ECEC samples from different cultures as well as partial gender invariance (Aboagye et al., 2019; Settanni et al., 2015; Tsigilis & Gregoriadis, 2008). However, it has been noted that professional caregivers report greater closeness in their relationship with girls than with boys (Howes et al., 2000; Solheim et al., 2012). Moreover, two items related to the *closeness* factor have been reported to be the main sources of model mis-specification in a Norwegian ECEC sample, namely, "this child spontaneously shares information about himself/herself" and "this child openly shares his/her feelings and experience with me" (Solheim et al., 2012). Additionally, the *closeness* factor has been shown to display minor in-variance problems across child genders for both an adaption of the full version of the STRS and the short version (Aboagye et al., 2019; Koomen et al., 2012), indicating that professional caregivers may, to some degree, perceive closeness in the relationship differently for girls and boys. For instance, the item "I share an affectionate, warm relationship with this child" has been shown to be

non-invariant across both child gender and age in a Ghanaian ECEC context (Aboagye et al., 2019), while the items "If upset, this child will seek comfort from me" and "this child is uncomfortable with physical affection or touch from me" were shown to be non-invariant across child gender in a Dutch sample (Koomen et al., 2012). Invariance issues have also been reported regarding children's age using the STRS when comparing ECEC and elementary school children (Koomen et al., 2012), suggesting that the factors *of closeness* and *conflict* may operate differently for younger and older children.

In Norway, 93.4% of children aged 1–5 years attend ECEC centers, where most spend 41 hr or more each week (Statistics Norway, 2022). This means that young children spend a considerable amount of time with professional caregivers. Even though there is some documentation of the psychometric properties of the original STRS in a Norwegian ECEC context based on a sample of Norwegian 4-year-old children (Solheim et al., 2012), the psychometric properties of the STRS-SF have not yet been investigated in Norway or in any other Nordic countries. Additionally, most studies that have investigated the psychometric properties of the STRS-SF in an ECEC context have only included children in the upper age range or used samples containing both ECEC and elementary school children. As the bidirectional professional caregiver-child relationship may vary across age groups depending on children's developmental maturity (Pianta et al., 2003), it is important to investigate the appropriateness of the STRS-SF for the youngest children, which is currently largely unknown. Hence, the aim of this study was to investigate and evaluate the factorial validity and measurement invariance of the STRS-SF in a sample applying the full age range of ECEC children aged 1–6 years old. As measurement invariance is a prerequisite for making meaningful comparisons between groups, such as children's age and gender, knowledge of this psychometric property is important for assessing the applicability of the STRS-SF.

Methods

This study is based on combined baseline data from two different ECEC projects conducted in central and south-eastern Norway. Baseline data from the project *Children in Central Norway* were collected over the period 2012–2014, whereas data from *Thrive by Three* were collected in 2018.

Procedure and Participants

Children in Central Norway. Parents of children enrolled in ECEC centers in three municipalities in central Norway received information letters regarding the project and via parent meetings before the project commenced. The information letter also provided the parents with the option to enroll their child into the project either by returning a signed consent form to their ECEC center or consenting digitally with their unique invitation code to the project's online survey. Parental consent was obtained to allow the professional caregiver who knew the child best to complete an online survey regarding the child. Professional caregivers provided written consent with their own unique invitation codes for the online survey. A total of 1631 parents (77%) of the invited parents consented to enroll their child in the project, and 169 professional caregivers reported on 1430 children between 1 and 6 years old (mAge = 44 months, 51% boys).

Thrive by Three. The baseline data from *Thrive by Three* included 1471 children (mAge = 21 months, 51% boys) and 184 units/groups from 78 ECEC centers. A professional caregiver within the unit/group who knew the child best answered the STRS-SF. When combining *Thrive by Three* with the sample from *Children in Central Norway*, the total sample for this study included 2901

children (mAge = 33 months, 51% boys) and 353 professional caregivers. On average, professional caregivers reported on 8.2 children each.

For both above-mentioned projects, participation was voluntary, and consent could be withdrawn without reprisal at any time until the participation registry was anonymized.

Measurement

The Student-Teacher Relationship Scale- Short Form.   As previously mentioned, the STRS-SF (Pianta, 2001) is a self-reported instrument regarding professional caregivers' perceptions of the relationship quality with a specific child. The instrument comprises 15 items (see Appendix for overview) measuring the two factors, *closeness* (eight items; e.g., "If upset, this child will seek comfort from me") and *conflict* (seven items; e.g., "This child and I always seem to struggle with each other") on a five-point Likert scale with response options ranging from 1 = Definitely does not apply to 5 = Definitely does apply. Thus, the *closeness* factor score ranged from 8 to 40, while the *conflict* factor score ranged from 7 to 35. Higher scores on the *closeness* factor indicate more positive interactions, while higher scores on the *conflict* factor indicate more negative interactions.

Statistical Analyses

Children's age was dichotomized into younger ECEC children (under 36 months) and older ECEC children (36 months and older) to reflect the organizational structure of ECEC centers in Norway. First, the internal consistency of the STRS-SF factors, *closeness* and *conflict*, for the full sample was investigated to obtain the multi-level omega (ω) coefficient. The ω coefficient was preferred over the more commonly used Cronbach's alpha because the latter depends on rather strict assumptions, such as tau-equivalence and normally distributed scores, which can lead to biased estimates if violated (Dunn et al., 2014; McNeish, 2018; Peters, 2014; Sijtsma, 2009; Yang & Green, 2011). The multi-level ω with 95% confidence interval was computed with the package "multilevelTools" (Wiley, 2022) in Rstudio. The multi-level ω estimates are interpreted in the same way as the alpha, where estimates ≥.70 are considered to indicate satisfactory internal consistency (Taber, 2018).

Before proceeding to the next step, the intraclass correlation (ICC) for the two factors was investigated separately and combined using STATA17 due to the nested structure of the data, as professional caregivers reported on average for 8.2 children each. The ICC was .23 for *closeness*, .20 for *conflict*, and the residual ICC was .23. Consequently, multi-level analyses were performed.
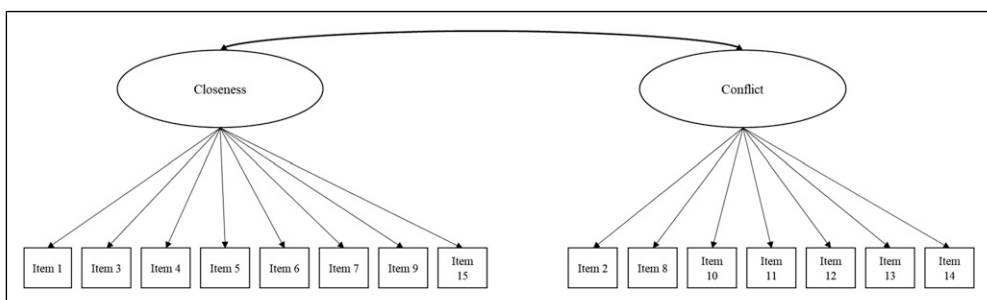

Figure 1. Test model for multi-group factor analyses based on children's gender and age. *Note.* Item overview can be found in the Appendix.

Next, we investigated the factorial validity and measurement invariance of the STRS-SF using a series of multi-level multi-group confirmatory factor analyses (MGCFA) based on children's age and gender (Figure 1), as this two-factor structure has shown promising psychometric properties in previous studies (Aboagye et al., 2019; Settanni et al., 2015; Tsigilis & Gregoriadis, 2008). The purpose of carrying out MGCFA is to determine whether the respondents attribute the same meaning to the latent factors as well as whether the means and scores can be interpreted similarly across groups (van de Schoot et al., 2012). This is done by investigating the model fit indices while adding additional constraints to the models following a hierarchical structure ranging from configural (weak invariance) to scalar (strong invariance), more specifically:

1. Configural invariance (equal factor structure across groups)
2. Metric invariance (equal factor loadings across groups)
3. Scalar invariance (equal thresholds across groups as the variables are ordered categorical).

Step 1 was to test the STRS-SF two-factor baseline model across children's age and gender (configural variance), where all parameters could vary freely. Step 2 was to test a model in which only the factor loadings were constrained between groups while the thresholds could vary freely (metric variance). In Step 3 we tested a model in which both the loadings and thresholds were constrained to be equal between groups (scalar invariance). Configural invariance exists if the two-factor model shows a good fit across the groups tested. Metric invariance exists if the more constrained model still shows a good model fit compared with the baseline model, whereas scalar invariance exists if the even more constrained model still shows a good model fit compared with the metric invariance model (Hirschfeld & Von Brachel, 2014).

The model fit was evaluated by inspecting the root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker-Lewis index (TLI). RMSEA values of $\leq .05$ indicate a good fit and values between .05 and .10 indicate an acceptable fit (MacCallum et al., 1996). For the CFI and TLI, values of $\geq .95$ are commonly used to indicate a good model fit (Hu & Bentler, 1999), however, Browne and Cudeck (1993) argue that these thresholds are too strict and rather recommend a threshold of $> .90$ to indicate a good model fit and values of .80–.90 indicate an acceptable model fit. For the evaluation of invariance, Cheung and Rensvold (2002) recommend that a CFI reduction of $\leq .01$ when adding additional constraints to the model indicates that the null hypothesis of invariance should not be rejected. The CFI difference between models was preferred as an indicator of invariance, as it is less sensitive to sample size and more sensitive to lack of variance than chi-square (χ^2) statistics (Meade et al., 2008). Multi-level MGCFA analyses were performed with Mplus v.8.4 (Muthén & Muthén, 1998-2017) using the weighted least square mean variance (WLSMV) estimator. The WLSMV estimator is appropriate for ordered categorical data and produces accurate parameter estimates (DiStefano & Morgan, 2014). Lastly, if scalar invariance was not found, we inspected the modification indices (χ^2) to locate non-invariant items and then relaxed the constraints for the non-variant items one by one, starting with the item with the greatest expected parameter change (EPC), to see if this improved the model fit. If the less constrained scalar invariant model showed a CFI estimate within the threshold of $\leq .01$, compared to the metric model, partial scalar invariance was observed.

Results

One child was excluded because of missing STRS-SF data, resulting in a final sample of 2900 children. The means and standard deviations for the STRS-SF items, as well as the *closeness* and *conflict* factors separated by children's gender and age are shown in Table 1. Regarding the

Table 1. Means and Standard Deviations of the STRS-SF Items and Scales for Boys and Girls Grouped by Age (N = 2900).

	Boys 1–2 years old (n = 972)	Boys 3–6 years old (n = 500)	Girls 1–2 years old (n = 944)	Girls 3–6 years old (n = 484)
Item 1	4.59 (.63)	4.27 (.65)	4.59 (.62)	4.33 (.69)
Item 2	1.30 (.78)	1.33 (.69)	1.23 (.68)	1.20 (.53)
Item 3	4.51 (.71)	4.14 (.80)	4.53 (.68)	4.24 (.79)
Item 4	4.56 (.96)	4.25 (1.02)	4.62 (.91)	4.53 (.88)
Item 5	4.70 (.60)	4.30 (.63)	4.66 (.60)	4.43 (.65)
Item 6	4.61 (.66)	4.43 (.68)	4.62 (.64)	4.54 (.66)
Item 7	3.17 (1.36)	4.02 (1.08)	3.32 (1.31)	4.19 (1.02)
Item 8	1.55 (.90)	1.39 (.73)	1.56 (.93)	1.28 (.63)
Item 9	4.25 (.89)	3.86 (1.04)	4.31 (.89)	3.96 (1.01)
Item 10	2.20 (1.20)	1.92 (1.20)	2.15 (1.19)	1.77 (1.10)
Item 11	1.32 (.77)	1.37 (.78)	1.26 (.69)	1.30 (.72)
Item 12	1.37 (.77)	1.44 (.83)	1.39 (.78)	1.42 (.75)
Item 13	1.44 (.85)	1.34 (.76)	1.44 (.83)	1.31 (.71)
Item 14	1.23 (.64)	1.23 (.61)	1.25 (.65)	1.30 (.69)
Item 15	4.03 (.96)	3.79 (1.10)	4.09 (.91)	4.01 (.98)
Closeness	34.42 (4.23)	33.06 (4.51)	34.74 (4.05)	34.24 (4.40)
Conflict	10.41 (3.78)	10.02 (4.07)	10.28 (3.76)	9.58 (3.78)

Note. Item overview can be found in the Appendix. Min–Max items = 1–5, Min–Max Closeness = 8–40, Min–Max Conflict = 7–35.

Table 2. STRS-SF Baseline (Configural) Model Fit Indices Grouped by Children's Gender and Age.

	CFI	TLI	RMSEA
Children's gender	.932	.920	.069
Children's age	.920	.906	.069

Note. CFI = comparative fit index, TLI = Tucker-Lewis index, RMSEA = root mean square error of approximation.

internal consistency of the factors for the full sample, the multi-level ω coefficient was .73 [CI .72, .75] for *closeness* within professional caregivers and .77 [CI .72, .83] between, while the *conflict* factor showed an estimate of .77 [CI .75, .78] within professional caregivers and .88 [CI .85, .92] between, indicating satisfactory internal consistency for both factors.

The fit indices for the unconstrained two-factor baseline model are shown in Table 2, indicating a good to acceptable overall model fit for the STRS-SF two-factor structure across children's gender (CFI = .932, TLI = .920, RMSEA = .069) and age (CFI = .920, TLI = .906, RMSEA = .069). Inspecting the CFI estimates in Table 3, the non-essential drop (≤.01) in CFI when adding additional constraints for children's gender provides support for both metric and scalar invariance for these groups. Regarding measurement invariance based on children's age, Table 3 shows a non-essential drop when constraining factor loadings but an essential drop (>.01) when constraints are added to the thresholds. Thus, metric invariance for age groups is supported, whereas scalar invariance is not.

As the STRS-SF showed scalar non-invariance based on children's age, modification indices were inspected. Following this, the threshold constraints for Item 7 ("This child spontaneously shares information about himself/herself," $\chi2$ = 160.09, EPC = −.42 for the younger age group

Table 3. STRS-SF Measurement Invariance Indices Using Comparative Fit Index (CFI) Estimates Grouped by Children's Gender and Age.

	Configural	Metric	Scalar
Children's gender	.932	.935	.941
Children's age	.920	.921	.904

and $\chi 2 = 160.12$, EPC = .42 for the older age group) related to the *closeness* factor, were relaxed. Relaxing the constraints for this item resulted in an increase in the CFI estimate from .904 to .924, which is higher than the >.01 threshold compared to the metric invariant model with a CFI of .921. Consequently, partial scalar invariance for the STRS-SF based on age group was evidenced when relaxing the constraints for item 7.

Discussion

This study aimed to investigate and evaluate the factorial validity of the STRS-SF (Pianta, 2001) and its measurement invariance across ECEC children's gender and age. As the instrument is frequently used to measure professional caregivers' perceptions of the relationship quality with a specific child, knowledge about how the instrument works within the ECEC context and across subgroups is pivotal to make accurate estimates and interpretations. This study is the first to investigate the psychometric properties of the STRS-SF in a Nordic ECEC context and expands the knowledge about the instrument's applicability globally by including the full age range of ECEC children, where children under the age of three have previously received little or no attention. Our findings indicate that the STRS-SF has promising psychometric properties regarding internal consistency, factorial validity, and measurement invariance. However, some precautions are warranted with respect to measurement invariance and need to be considered.

Factorial Validity and Measurement Invariance

In line with previous studies (Aboagye et al., 2019; Settanni et al., 2015; Tsigilis & Gregoriadis, 2008), this study found promising psychometric properties of the STRS-SF. The instrument showed satisfactory internal consistency, a good to acceptable model fit for the two-factor structure, and full or partial measurement invariance across children's gender and age. Similar to previous studies investigating the measurement invariance of the STRS-SF, some items are flagged as non-invariant. Our findings showed full scalar invariance across children's gender, indicating that the latent means of the *closeness* and *conflict* factors can be compared meaningfully between boys and girls, as they show similar structure and meaning across the groups. In other words, professional caregivers interpreted the items similarly, regardless of the child's gender. Regarding the children's age, only partial scalar invariance was found. This indicates that the factor structure and strength of the factor loadings operate similarly between older and younger ECEC children, while caution is warranted when comparing latent means from the *closeness* factor, as one item operated differently depending on the child's age. Non-invariant items based on children's age related to the *closeness* factor have been reported in previous studies (Aboagye et al., 2019; Koomen et al., 2012), which is also the case in this study. However, and contrary to previous studies, the main sources for not finding full scalar invariance across children's age in this study were due to the item "this child spontaneously shares information about himself/herself". This indicates that professional caregivers interpret this item differently in older and younger ECEC children, which consequently may bias the *closeness* factor as it contains an item that holds

different meaning across age groups. Additionally, the same item was also flagged by Solheim et al. (2012) as one of the main sources of model misspecification in the original full version of the STRS among Norwegian four-year-old children.

As the above-mentioned non-invariant item has a communication element to it, it is plausible to think that older and younger ECEC children varies in their degree of spontaneity and ability to share information about themselves, which in turn will lead professional caregivers to interpret this item differently. In other words, as children matures and develop their cognitive, communicative, and self-regulatory capacities, their communicative expression changes, which in turn may influence the professional caregivers' interpretations of this item on the STRS-SF. This said, it has been shown that both professional caregivers and children characteristics influence professional caregivers' perception of the child-professional caregiver relationship (Berg-Nielsen et al., 2012; Choi & Dobbs-Oates, 2016). For instance, children with autism spectrum disorders have difficulties with initiating and maintaining social relationships (i.e., social communication) (Yoder et al., 2014) and it has been reported that children's level of autistic mannerisms are negatively related to professional caregivers' perception of closeness in the relationship (Blacher et al., 2014). Consequently, users of the STRS-SF should be aware of factors that may influence the ratings on the instrument.

Overall, this study supports the applicability of the STRS-SF for younger and older children in an ECEC context and across children's genders, even though the STRS-SF was originally intended to be used with children from the age of four. However, awareness of the non-variant item related to the *closeness* factor based on children's age is warranted when applying the instrument, as well as the rapid pace children develop during the preschool period and the limited capacities and behavioral repertoire for the youngest children. One way to deal with age-related non-invariance and developmental processes may be to use age-specific norms, acknowledging that even though the factor structure and factor loadings are comparable, the latent means may hold different meanings and thus may not be directly comparable.

Strengths and Limitations

The main strength of this study is the inclusion of two large community samples covering the full age range of children enrolled in ECEC centers. One limitation of the current study is the non-convergence of data when investigating the psychometric properties of the STRS-SF using one-year-intervals rather than dichotomizing children's age in under 36 months and 36 months and older. From a developmental perspective, future research should investigate this further to pinpoint more precisely at what age non-invariance is introduced to the instrument. However, in a Norwegian context this dichotomization of children's age reflects the organizational structure of ECEC centers in Norway, as children are grouped based on their age. Another possible limitation is the social desirability bias, which would be present if professional caregivers feel that forming high-quality relationships with children is expected, leading them to report higher relationship quality than it actually is. In this study, we instructed professional caregivers who knew the child best to complete the STRS-SF. Normally, children will interact with several adults during their days in an ECEC center. Unfortunately, an inter-rater approach was not available for this study. Thus, future research should investigate the inter-rater agreement between staff members in ECEC centers to more closely investigate how perceptions of relationship quality are formed and factors related to it. Additionally, investigating the psychometric properties of the STRS-SF using other subgroups should be explored further (e.g., children developing normally vs. developmentally delayed). Last, even though the findings from this study lend support to the factor structure of the STRS-SF, the non-invariant items seem to differ between cultures, indicating that professional caregivers from different cultures seem to interpret items differently. Consequently, an interesting

aim for future research would be to investigate the STRS-SF from a multicultural perspective to see how culture influences the perception of relationship quality.

Conclusion

This study adds to the knowledge about the psychometric properties of the STRS-SF, both globally and in Nordic countries, supporting its applicability in an ECEC context. The findings indicate that the STRS-SF can be indiscriminately applied to both younger and older children in ECEC centers based on children's gender, but some caution is warranted when comparing latent mean scores of older and younger children due to a non-invariant item related to the *closeness* factor, indicating that this item is interpreted differently by professional caregivers across children's age groups.

Appendix

Overview of the items included in STRS-SF and their related factor.

Closeness	1. I share an affectionate, warm relationship with this child
	3. If upset, this child will seek comfort from me
	4. This child is uncomfortable with physical affection or touch from me
	5. This child values his/her relationship with me
	6. When I praise this child, he/she beams with pride
	7. This child spontaneously shares information about himself/herself
	9. It is easy to be in tune with what this student is feeling
	15. This child openly shares his/her feelings and experience with me
Conflict	2. This child and I always seem to be struggling with each other
	8. This child easily becomes angry at me
	10. This child remains angry or is resistant after being disciplined
	11. Dealing with this child drains my energy
	12. When this child arrives in a bad mood, I know we're in for a long and difficult day
	13. This child's feelings toward me can be unpredictable or can change suddenly
	14. This child is sneaky or manipulative with me

Note. Item four is reversed before calculating the Closeness score.

ORCID iD

K. Stensen https://orcid.org/0000-0001-8040-9006

References

Aboagye, M. O., Qin, J., Pekárková, S., Antwi, C. O., Jababu, Y., Asare, K., Affum-Osei, E., & Akinyi, N. (2019). 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. *Psychological Studies*, *64*(2), 221–234. https://doi.org/10.1007/s12646-019-00488-0

Berg-Nielsen, T. S., Solheim, E., Belsky, J., & Wichstrøm, L. (2012). Preschoolers' psychosocial problems: In the eyes of the beholder? Adding teacher characteristics as determinants of discrepant parent-teacher reports. *Child Psychiatry and Human Development*, *43*(3), 393–413. https://doi.org/10.1007/s10578-011-0271-0

Beyazkurk, D., & Kesner, J. E. (2005). Teacher-child relationships in Turkish and United States schools: A cross-cultural study. *International Educational Research*, *6*(5), 547–554.

Blacher, J., Howell, E., Lauderdale-Littin, S., DiGennaro Reed, F. D., & Laugeson, E. A. (2014). Autism spectrum disorder and the student teacher relationship: A comparison study with peers with intellectual disability and typical development. *Research in Autism Spectrum Disorders*, *8*(3), 324–333. https://doi.org/10.1016/j.rasd.2013.12.008

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.

Center on the Developing Child at Harvard University (2010). The foundations of lifelong health are built in early childhood. Retrieved from http://www.developingchild.harvard.edu

Center on the Developing Child at Harvard University (2009). Young children develop in an environment of relationships. Retrieved from http://www.developingchild.harvard.edu

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

Choi, J. Y., & Dobbs-Oates, J. (2016). Teacher-child relationships: Contribution of teacher and child characteristics. *Journal of Research in Childhood Education*, *30*(1), 15–28. https://doi.org/10.1080/02568543.2015.1105331

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 425–438. https://doi.org/10.1080/10705511.2014.915373

Doumen, S., Verschueren, K., Buyse, E., De Munter, S., Max, K., & Moens, L. (2009). Further examination of the convergent and discriminant validity of the student-teacher relationship scale. *Infant and Child Development*, *18*(6), 502–520. https://doi.org/10.1002/icd.635

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

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. https://doi.org/10.1111/bjop.12046

Fraire, M., Longobardi, C., Prino, L. E., Sclavo, E., & Settanni, M. (2017). Examining the student-teacher relationship scale in the Italian context: A factorial validity study. *Electronic Journal of Research in Education Psychology*, *11*(31), 851–882. https://doi.org/10.25115/ejrep.31.1591

Graziano, P. A., Garb, L. R., Ros, R., Hart, K., & Garcia, A. (2016). Executive functioning and school readiness among preschoolers with externalizing problems: The moderating role of the student-teacher relationship. *Early Education and Development*, *27*(5), 573–589. https://doi.org/10.1080/10409289.2016.1102019

Griggs, M. S., Gagnon, S. G., Huelsman, T. J., Kidder-Ashley, P., & Ballard, M. (2009). Student-teacher relationships matter: Moderating influences between temperament and preschool social competence. *Psychology in the Schools*, *46*(6), 553–567. https://doi.org/10.1002/pits.20397

Hirschfeld, G., & Von Brachel, R. (2014). Improving multiple-group confirmatory factor analysis in R-A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research, and Evaluation*, *19*, 1–12. https://doi.org/10.7275/qazy-2946.

Howes, C., Phillipsen, L. C., & Peisner-Feinberg, E. (2000). The consistency of perceived teacher–child relationships between preschool and kindergarten. *Journal of School Psychology*, *38*(2), 113–132. https://doi.org/10.1016/s0022-4405(99)00044-8

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

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

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149. https://doi.org/10.1037/1082-989x.1.2.130

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433. https://doi.org/10.1037/met0000144

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *The Journal of Applied Psychology*, *93*(3), 568–592. https://doi.org/10.1037/0021-9010.93.3.568

Miller-Lewis, L. R., Sawyer, A. C. P., Searle, A. K., Mittinty, M. N., Sawyer, M. G., & Lynch, J. W. (2014). Student-teacher relationship trajectories and mental health problems in young children. *BMC Psychology*, *2*(1), 27. https://doi.org/10.1186/s40359-014-0027-2

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.). Muthén and Muthén.

OECD (2021). *OECD family database*. OECD Publishing. https://www.oecd.org/els/soc/PF3_2_Enrolment_childcare_preschool.pdf

Peters, G. J. (2014). The alpha and the omega of scale reliability and validity: Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *The European Health Psychologist*, *16*(2), 56–69. https://www.ehps.net/ehp/index.php/contents/article/view/ehp.v16.i2.p56.

Pianta, R. C. (2001). *STRS Student-Teacher Relationship Scale (professional manual)*. Psychological Assessment Resources, Inc.

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. 199–234). John Wiley and Sons Inc.

Sabol, T. J., & Pianta, R. C. (2012). Recent trends in research on teacher-child relationships. *Attachment and Human Development*, *14*(3), 213–231. https://doi.org/10.1080/14616734.2012.672262

Schmitt, M. B., Pentimonti, J. M., & Justice, L. M. (2012). Teacher-child relationships, behavior regulation, and language gain among at-risk preschoolers. *Journal of School Psychology*, *50*(5), 681–699. https://doi.org/10.1016/j.jsp.2012.04.003

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*, 898. https://doi.org/10.3389/fpsyg.2015.00898

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/s11336-008-9101-0

Skalická, V., Stenseng, F., & Wichstrøm, L. (2015). Reciprocal relations between student-teacher conflict, children's social skills and externalizing behavior: A three wave longitudinal study from preschool to third grade. *International Journal of Behavioral Development*, *39*(5), 413–425. https://doi.org/10.1177/0165025415584187

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

Statistics, N. (2022). Kindergartens. Retrieved from https://www.ssb.no/barnehager

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tsigilis, N., & Gregoriadis, A. (2008). Measuring teacher-child relationships in the Greek kindergarten setting: A validity study of the student-teacher relationship scale–short form. *Early Education and Development*, *19*(5), 816–835. https://doi.org/10.1080/10409280801975826

Tsigilis, N., Gregoriadis, A., & Grammatikopoulos, V. (2018). Evaluating the student-teacher relationship scale in the Greek educational setting: An item parcelling perspective. *Research Papers in Education*, *33*(4), 414–426. https://doi.org/10.1080/02671522.2017.1353675

van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486–492. https://doi.org/10.1080/17405629.2012.686740

Verschueren, K., & Koomen, H. (2021). Dependency in teacher-child relationships: Deepening our understanding of the construct. *Attachment and Human Development*, *23*(5), 481–489. https://doi.org/10.1080/14616734.2020.1751986

Verschueren, K., & Koomen, H. M. Y. (2012). Teacher-child relationships from an attachment perspective. *Attachment and Human Development*, *14*(3), 205–211. https://doi.org/10.1080/14616734.2012.672260

Wiley, J. F. (2022). Multilevel and mixed effects model diagnostics and effect sizes. Retrieved from https://cran.rproject.org/web/packages/multilevelTools/multilevelTools.pdf

Yang, Y., & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, *29*(4), 377–392. https://doi.org/10.1177/0734282911406668

Yoder, P. J., Bottema-Beutel, K., Woynaroski, T., Chandrasekhar, R., & Sandbank, M. (2013). Social communication intervention effects vary by dependent variable type in preschoolers with autism spectrum disorders. *Evidence-based Communication Assessment and Intervention*, *7*(4), 150–174. https://doi.org/10.1080/17489539.2014.917780

Zhan, X., & Sun, J. (2011). The reciprocal relations between teachers' perceptions of children's behavior problems and teacher-child relationships in the first preschool year. *The Journal of Genetic Psychology*, *172*(2), 176–198. https://doi.org/10.1080/00221325.2010.528077