



The Student–Teacher Relationship Scale revisited: Testing factorial structure, measurement invariance and validity criteria in German-speaking samples



Anne Milatz^{a,*}, Michael Glüer^{b,c}, Elena Harwardt-Heinecke^a,
Gregor Kappler^a, Lieselotte Ahnert^a

^a University of Vienna, Faculty of Psychology, Department of Applied Psychology: Development, Health, Intervention and Enhancement, Liebiggasse 5, 1010 Vienna, Austria

^b University of Bielefeld, Faculty of Psychology and Sport Science, Department of Psychology: Developmental Psychology, Universitätsstrasse 25, 33615 Bielefeld, Germany

^c Freie Universität Berlin, Department of Educational Science and Psychology, Habelschwerdter Allee 45, 14195 Berlin, Germany

ARTICLE INFO

Article history:

Received 15 February 2013

Received in revised form 21 March 2014

Accepted 3 April 2014

Available online 19 April 2014

Keywords:

German STRS

Factorial structure

Measurement invariance

Construct validity

Criterion validity

ABSTRACT

The Student–Teacher Relationship Scale (STRS) is widely used for research in kindergarten and school. The increasing number of applications inside and outside of the U.S. stresses the need to investigate STRS properties, accordingly. The present study used the STRS in German-speaking countries, examining whether (a) the original factor structure is appropriate for a German version, (b) whether applications of a German STRS are invariant across contexts (kindergarten, first and second grade) as well as gender, and (c) whether construct and criterion validity are met. The original STRS was translated into German and filled out by 368 kindergarten and 503 elementary school teachers in Germany and Austria. Observations in kindergartens, student reports in schools, and teacher reports of students' characteristics served as validity criteria. Results of confirmatory factor analyses (CFAs) did not confirm the original STRS factor structure. Subsequent exploratory factor analyses on training samples resulted in significant item reductions, followed by further CFAs on validation samples. The bootstrapped results yielded an adjusted three-factor model with subscales indicating satisfying alphas and invariance across context and gender. Construct and criterion validity were met for all subscales of the German STRS based on various criteria from both, observations and reports.

© 2014 Elsevier Inc. All rights reserved.

The Student–Teacher Relationship Scale (STRS; Pianta, 2001) has been widely used to assess student–teacher relationship qualities in kindergarten and school, as well as to explore the impact of student–teacher relationships on a large variety of social, emotional (Baker, 2006; Birch & Ladd, 1998; Hamre & Pianta, 2001) and academic (Birch & Ladd, 1997; Hughes & Cavell, 1999; Palermo, Hanish, Martin, Fabes, & Reiser, 2007) outcomes in students (Davis, 2003; Roorda, Koomen, Spilt, & Oort, 2011). This empirical research has emerged over the last 20 years (Bretherton, 1992; Pianta, 1999; Pianta, Hamre, & Stuhlman, 2003) extending the traditional focus in attachment research on mother–child relationships to

relationships of students toward their teachers (Hamilton & Howes, 1992; Howes & Matheson, 1992; Pianta, 1992). Whereas earlier studies within this research field linked student–teacher relationships to students' later development, later investigations explored kindergarten and school contexts (Davis, 2003; Pianta & Nimetz, 1991; Pianta & Steinberg, 1992; Pianta, Steinberg, & Rollins, 1995) and provided more insight into how children develop relationships toward their teachers differently from those to their parents (Ahnert, Piquart, & Lamb, 2006).

The items of the STRS capture the student–teacher relationship quality (Pianta, 2001) with the use of three subscales, namely, closeness, conflict and dependency. Closeness refers to warmth and affection, conflict characterizes negativity and unpredictability and dependency concerns a student's degree of autonomy (reversed rated) within the relationship. Closeness has been shown to have positive associations with school adjustment (Arbeau, Coplan, & Weeks, 2010; Baker, 2006; Buyse, Verschueren, Verachtert, & van Damme, 2009; Pianta et al., 1995), prosocial behaviors

* Corresponding author. Tel.: +43 1 4277 47227; fax: +43 1 4277 847227.

E-mail addresses: anne.milatz@univie.ac.at (A. Milatz), michael.gluer@uni-bielefeld.de (M. Glüer), elena.harwardt@univie.ac.at (E. Harwardt-Heinecke), gregor.kappler@univie.ac.at (G. Kappler), lieselotte.ahnert@univie.ac.at (L. Ahnert).

(Hughes & Cavell, 1999; Palermo et al., 2007), effectiveness of task performance (Ahnert, Milatz, Kappler, Schneiderwind, & Fischer, 2013), levels of academic performance (Birch & Ladd, 1997; Peisner-Feinberg et al., 2001; Spilt, Hughes, Wu, & Kwok, 2012) as well as stress-regulation (Ahnert, Harwardt-Heinecke, Kappler, Eckstein-Madry, & Milatz, 2012). In contrast, conflict has been associated with social withdrawal, antisocial behaviors, school maladjustment, and deficits in math and language skills (Birch & Ladd, 1998; Buyse et al., 2009; Hamre & Pianta, 2001; Murray, Murray, & Waas, 2008; Palermo et al., 2007; Rudasill & Rimm-Kaufman, 2009). Dependency has been positively linked to behavior problems (Pianta, 2001) and there have been negative associations with school adjustment and mastery motivation (Birch & Ladd, 1997; Birch & Ladd, 1998; Hamre & Pianta, 2001; Pianta, 2001).

Overall, good student–teacher relationships facilitate learning by providing emotional support and assistance for students' exploration (Booth, Kelly, Spieker, & Zuckerman, 2003; Hamilton & Howes, 1992; Pederson & Moran, 1995). Good relationships are, in general, characterized by closeness, by which the student accepts the teacher as an important emotional and cognitive resource (Mashburn & Pianta, 2006; Pianta, Nimetz, & Bennett, 1997; Pianta, 1992). Thus, the teacher serves as a secure-base which is a core element within the attachment theory framework. Although the secure base concept was originally thought of as universal in social relations, which suspends cross-cultural differences in establishing close relationships (Posada et al., 1995; van Ijzendoorn & Sagi, 2008), recent research in cross-cultural psychology has been discovering different cultural pathways, at least for mother-child relationships. For example, cultural contexts focusing on the independence of individuals appreciate autonomy and limit the formation and maintenance of relationships, whereas cultural contexts focusing on interdependence are oriented toward relatedness and thus promote relationships among individuals, especially if commitment and obedience receive high social value (Greenfield, Keller, Fuligni, & Maynard, 2003). Thus, these cultural influences on relationships might be also discovered when applying the STRS in different countries. Not surprisingly, closeness and conflict of the STRS correlated negatively across applications in the US, Greece, Norway and the Netherlands (ranging from $-.34$ to $-.62$), even though dependency revealed cultural differences. In Greece, for example, dependency was positively associated with closeness ($r = .35$) and was unrelated to conflict (Gregoriadis & Tsigilis, 2008) which stands in contrast to the American studies (Pianta, 2001; Webb & Neuharth-Pritchett, 2011) in which dependency was positively related to conflict ($r = .28$) and barely associated with closeness ($r = .13$). This weak correlation is similar to the Dutch study which showed a correlation of closeness and dependency of $r = .05$. Thus, dependency might be viewed as something desirable in Greece, whereas in the US and the Netherlands, a student's salient need for interaction with their teachers might be perceived as inappropriate and even aggressive (Coplan & Prakash, 2003; Sroufe, 2005; Sroufe, Fox, & Pancake, 1983).

Because the subscales of the STRS seem to have been perceived differently across western countries where cross-cultural differences are rather small, the present paper aims to investigate how far a German application of the STRS and its subscales resemble or deviate the original and other language versions. Regarding the educational systems in Europe, it is important to note, that differences in some cases are considerable. For instance, students in the Netherlands stay in basic school from age four to eight, whereas the German speaking students stay in kindergarten until the age of six and enter primary school thereafter. The German speaking kindergarten setting is less structured than the German speaking school setting and thus comprises different relationship structures between students and teachers whose understanding of their roles and functions differ. Whereas kindergarten teachers are focused on

providing emotional climates of acting and learning, elementary school teachers see themselves more as instructors who provide cognitive challenges and didactic elaborated teaching.

Surprisingly few European studies have explored whether the factorial structure of the STRS remained the same when used in the respective countries, applying the STRS in their native language. Italian and Greek applications of the STRS in kindergarten revealed the original three factor structure (Fraire, Longobardi, & Sclavo, 2008; Gregoriadis & Tsigilis, 2008) based on Principal Component Analysis (PCA).

Recently and independent from our work presented here, a German kindergarten teacher–child relationship measure was proposed in a study of Mayr (2012) including a translation of the original STRS and 17 further items (five items from the Parent–Child Reunion Inventory of Niederhofer (2000) and 12 self-developed items). Mayr's scale is hardly comparable with the original STRS by Pianta (2001), above all because Mayr's measure was specifically developed to capture the kindergarten context which is reflected in his translations of the STRS items and the new items included. A German translation capturing both, the kindergarten and school context, for which the original STRS was designed, is still lacking today. Especially for longitudinal research, such a measurement tool is needed. Furthermore, no validity criteria or psychometric analysis, such as invariance tests across these contexts were presented in Mayr's work. Also, one might criticize methodological issues, as missing data were substituted by mean values and the analysis of the factor structure of the German childcare teacher–child relationship measure (Mayr, 2012) has been based on PCA statistics only.

Whereas PCAs, however, do not reveal the underlying latent factor structure of an item set (Widaman, 1993), an Exploratory Factor Analysis (EFA) is considered to explore the factorial structure appropriately, and a Confirmatory Factor Analysis (CFA), examines a proposed factorial structure on its empirical foundation (Bryant & Yarnold, 1995). When the CFA framework was applied for the STRS recently used in American and Dutch kindergartens, as well as in Dutch and Norwegian schools, the original item set had to be reduced (Koomen, Verschueren, van Schooten, Jak, & Pianta, 2012; Webb & Neuharth-Pritchett, 2011). Due to model fit problems in Norway (Drugli & Hjemdal, 2013), a CFA even supported the short original STRS version with 15 items, involving the two subscales: Conflict and closeness (see also Tsigilis & Gregoriadis, 2008). These short versions, however, had been easily accepted (Baker, 2006; Howes, Hamilton, & Philipsen, 1998; O'Connor & McCartney, 2006; Spilt et al., 2012) as dependency had been frequently criticized due to its low alpha ($\alpha = .64$; Pianta, 2001), and less often exploited than closeness and conflict, whose alphas ranged between $.86$ and $.92$.

Growing research on different trajectories for children's gender, age, or ethnicity raises the question as to whether STRS applications across these variations are comparable (Buyse et al., 2009; Spilt et al., 2012). Above all, researchers in early childhood research applying the STRS are interested in longitudinal learning processes. Thus, a measurement tool whose factors can be meaningfully interpreted over time is of great importance (Birch & Ladd, 1998; Hamre & Pianta, 2001; Howes, 2000; Ladd & Price, 1987; Pianta & Stuhlman, 2004). However, measurement invariance, which assures that comparable factors among different groups or time points are captured (Meredith, 1993; Millsap, 2011), can be tested by including stepwise equality constraints to the model: Configural invariance is given when the same measurement model holds true across groups; weak factorial invariance when factor loadings are also equal; strong factorial invariance when additional intercepts are equal across groups; and strict factorial invariance can be assumed when residual variance is held equal across groups (Widaman & Reise, 1997).

Up until now, only a few researchers have examined invariance when applying the STRS. The STRS scales were shown not to be invariant regarding ethnicity in a U.S. sample involving African Americans and Caucasians (Webb & Neuharth-Pritchett, 2011) whereas invariance for gender was revealed in the Netherlands and Greece (Koomen et al., 2012; Tsigilis & Gregoriadis, 2008). Regarding age, weak measurement invariance for children from 3 to 8 years could be found (Koomen et al., 2012). However, it is still unclear as to whether the STRS might capture similar factors in kindergarten and elementary school, especially due to the fact that the two contexts are fairly contrasting. Thus, the present paper will address this issue because kindergarten and elementary school in the German speaking countries have somewhat different learning environments, varying in degrees of structured learning activities and of individualized interactions with teachers (Glüer, 2012).

In general, validation is needed to justify the interpretation of the STRS, specifically when the STRS has been used in different contexts outside of the U.S. where it was originally developed (Peng, Nisbett, & Wong, 1997; Saal, Downey, & Lahey, 1980). In several U.S. and a few European studies, the STRS was certified for criterion validity which links the STRS scales reasonably to child outcomes, as well as for construct validity which links the STRS scales to similar external measures on closeness, conflict, and dependency. In recent research, school adjustment, behavior problems, peer acceptance and task performance confirmed criterion validity (Ahnert et al., 2013; Arbeau et al., 2010; Baker, 2006; Birch & Ladd, 1997; Birch & Ladd, 1998; Buyse et al., 2009; Hamre & Pianta, 2001; Hughes & Cavell, 1999; Murray et al., 2008; Palermo et al., 2007; Peisner-Feinberg et al., 2001; Pianta, 1994; Pianta et al., 1995). Furthermore, child reports (Howes et al., 1998; Mantzicopoulos & Neuharth-Pritchett, 2003; Rey, Smith, Yoon, Somers, & Barnett, 2007; Valeski & Stipek, 2001), child drawings (Harrison, Clarke, & Ungerer, 2007), as well as observations by external observers on the relationship (Doumen et al., 2009; Howes & Ritchie, 1999; O'Connor & McCartney, 2006) have met the requirements for construct validity.

In this paper we present a new German STRS translation and several analyses showing the validity of this questionnaire ("German" refers to the language per se and the two populations Germany and Austria in which app. 80% of the German speaking population live). For construct validity, we chose an attachment measure (external observation) and a relationship measure (child report) which assesses the teacher's popularity. Specifically for closeness of the student-teacher relationships, which has its theoretical background in the attachment framework, positive associations between attachment security and warm affection toward the teacher were expected. Conflicts in the student-teacher relationships, however, might be reflected in impairments of the relationship, such as attachment insecurity as well as denials regarding the teacher. In addition, dependency of the student-teacher relationship might be linked to attachment insecurity, because dependency (or a lack of autonomy) has been linked to child behaviors such as displaying demanding prompts of attention, acts of impulsivity and less behavioral control (Kennedy & Kennedy, 2004). Regarding criterion validity, we used measures regularly applied for this purpose in the field, capturing prosocial behaviors and interpersonal competencies which are expected to be positively linked with closeness, but negatively connected with conflict and dependency. In contrast, children's behavioral and emotional difficulties might be negatively associated with closeness and positively with conflict and dependency. Moreover, negative attitudes toward school and learning are expected to be positively linked with conflict and dependency.

In sum, the outlined literature shows that the STRS is widely used, and that more analyses were repeatedly requested in order to shed further light on the psychometric properties of the STRS,

such as its factorial structure, invariance and validity in different languages and schooling cultures. Moreover, there exists currently no STRS measure which is suitable to assess student-teacher relationship quality in both kindergarten and school contexts in German-speaking countries.

The present study aims to substantially contribute to further understand the psychometric characteristics of the STRS within an international perspective, by presenting a new German translation of the STRS that is suitable for both the kindergarten and school context in German-speaking countries. Therefore the present study will examine (a) whether the original factor structure is appropriate for the German STRS (b) whether the application of the German STRS is invariant across contexts (kindergarten and school) and gender and (c) whether construct and criterion validity are met.

1. Method

1.1. Recruitment

Students were merged together from four different research projects in kindergartens and elementary schools in Germany and Austria dealing with student-teacher relationships and their impact on selected developmental domains. The research projects were carried out in large cities (Berlin and Cologne in Germany and Vienna in Austria) in middle-sized towns (such as Stendal in Germany and St. Poelten in Lower Austria) as well as their rural surroundings. Fifty-eight percent of the students lived in metropolises (e.g., Berlin, Cologne or Vienna), 29% were in medium-sized towns (e.g. Stendal or St. Poelten) and 13% in rural areas.

Participants were recruited using telephone lists of kindergartens and schools. Of the families approached, 50% agreed and those did not differ in terms of SES from the rest. We then selected approximately ten percent of the families balanced across gender, SES, and districts. As a result, sample composition varied over all strata of the population in Germany and Austria. However, various student-teacher and teacher-institution ratios appeared due to the research aims of the different studies (see sample composition Table 1).

1.2. Participants

1.2.1. Students and families

A total of $N=871$ students (451 girls) were enrolled in the present study, of which $n=368$ students (177 boys) attended kindergarten (mean age = 6;08 years, $SD=0;5$), $n=411$ students (198 boys) attended first grade (mean age = 7;1, $SD=0;8$) and 92 students (45 boys) were in second grade of elementary school (mean age = 7;9 years, $SD=0;3$). None of the children in the sample had repeated a class. The majority of the children (81.3%) lived with both parents and 18.7% only with their mothers. Parents widely represented middle-class families, with 46.3%/47.1% of the mothers/fathers having graduated from university, 51.7%/48.9% of them had finished vocational-technical school, 1.5%/2.3% had graduated from high school and only .5%/1.7% had dropped out of school. Household income per month was greater than 5000 € for 6.5% of the families (to convert € in US\$ multiply with app. 1.4, status 31.01.2014). It ranged between 3400 and 5000 € for 28.8%, 1200 and 3400 € for 57.9%, and was lower than 1200 € for only 6.8% of the families, reflecting SES distribution nationwide in Germany and Austria (Statistical Federal Office Destatis, 2011; Statistik Austria, 2012). Of the total student sample, the vast majority were Caucasians (96.7% in total, 95.4% of the kindergarteners and 97.1% of the school children) and the others had a migration background (from Pakistan, Turkey, Spain, Portugal and Italy). Thus culturally diverse children were underrepresented

Table 1
Sample composition across the four different research projects.

Sample characteristics	Berlin districts			Cologne districts		Stendal and surroundings		Vienna, St. Poelten and surroundings	
	Kiga	1st grade	2nd grade	Kiga	1st grade	Kiga	1st grade	Kiga	1st grade
No. of students	130	95	92	101	89	76	176	61	51
No. of girls	66	50	47	53	47	39	89	33	27
No. of teachers	130	91	89	56	57	33	66	18	8
No. of institutions	112	80	74	28	39	12	9	12	3
Student–teacher ratio	1	1.04	1.14	1.8	1.56	2.30	2.67	3.39	6.38
Teacher–institution ratio	1.16	1.14	1.20	2	1.46	2.75	7.3	1.5	2.67

Note. Kiga = kindergarten.

compared with the population (app. 20% immigrants; [Statistical Federal Office Destatis, 2012](#); [Statistik Austria, 2013](#)).

1.2.2. Teachers and institutions

Throughout the four research projects, $n=237$ kindergarten teachers from 164 kindergartens and $n=230$ elementary school teachers from 132 elementary schools participated (for more detail see [Table 1](#)). Each teacher had known the target students for a minimum of six months. The number of students in the classrooms at the kindergartens averaged 19.6 ($SD=3.6$) and 21.3 ($SD=4.1$) at elementary school. Teachers' ages in kindergarten ranged from 20 to 58 years ($M=40;4$, $SD=9;8$) and from 23 to 61 years ($M=42;1$, $SD=10;1$) for elementary school teachers, with professional experience of averaging 18;7 ($SD=10;0$) years in kindergarten and 19;6 years ($SD=11;5$) in elementary school, respectively.

1.3. Study design

All participating teachers filled out the Student–Teacher Relationships Scale (STRS; [Pianta, 2001](#)) for the entire sample of students. Further measures were applied either in kindergarten or at school or in both (depending on the type of measure), which later served as validity criteria in the present study. In kindergarten, external observers evaluated the quality of teacher–student attachments on 48% ($n=177$) of the kindergarten sample using the Attachment-Q-Sort (AQS; [Waters, 1995](#)). In kindergarten and at school, teachers filled out the Interpersonal Competence Scale (ICS; [Cairns, Leung, Gest, & Cairns, 1995](#)) for 49% ($n=179$) of the kindergarten and for 37% ($n=187$) of the school sample, as well as the Strengths and Difficulty Questionnaire (SDQ; [Klasen et al., 2000](#); [Woerner, Becker, & Rothenberger, 2004](#)) for 17% ($n=61$) of the kindergarten and for 18% ($n=89$) of the school sample. In elementary school, 18% ($n=89$) of the students were interviewed using Students' Report on Feelings about School ([Rauer & Schuck, 2004](#)).

Because three of the applied measures (STRS, ICS, AQS) were available only in English, two independent researchers with expertise in attachment theory and relationship issues conducted the parallel blind technique according to translation standards ([Behling & Law, 2005](#)), whereupon one of them was a bilingual expert and one proved English skills at a professional level.

1.4. Measures

1.4.1. Student–Teacher Relationships Scale

The STRS provides three subscales, comprising 28 items in total. Conflict (12 items) describes the extent to which teachers evaluate the relationship as negative or unpredictable as well as how emotionally drained they feel by a student. Closeness (11 items) rates how warm, affectionate, and supportive the teacher evaluates himself to be for a student. Dependency (5 items) measures the extent to which a teacher evaluates a student as overly dependent, resistant, and clingy. The five-point likert items range from 1 (“definitely does not apply”) to 5 (“definitely applies”). Over the present

data set, Cronbach's alphas for closeness were $\alpha_{\text{kindergarten}} = .80$ and $\alpha_{\text{school}} = .77$, for conflict $\alpha_{\text{kindergarten}} = .66$ and $\alpha_{\text{school}} = .72$ and for dependency $\alpha_{\text{kindergarten}} = .49$ and $\alpha_{\text{school}} = .54$, respectively.

1.4.2. Interpersonal Competence Scale

The ICS aims to capture students' social and behavioral characteristics on three subscales: (a) aggression (3 items: $\alpha_{\text{kindergarten}} = .61$, $\alpha_{\text{school}} = .67$) assesses the extent to which a student argues, troubles and fights; (b) popularity (3 items: $\alpha_{\text{kindergarten}} = .65$, $\alpha_{\text{school}} = .69$) is concerned with how popular a student among boys, girls, and friends in general; and (c) social affiliation (2 items: $\alpha_{\text{kindergarten}} = .68$, $\alpha_{\text{school}} = .65$) describes the extent to which a student is friendly. Items are rated on a seven-point likert scale.

1.4.3. Strength and Difficulty Questionnaire

The 25 items of an widely used German version of the SDQ ([Klasen et al., 2000](#)) determine five subscales listing students' emotions and behaviors as follows: (a) emotional symptoms (e.g. fears, worries, clingy: $\alpha_{\text{kindergarten}} = .62$, $\alpha_{\text{school}} = .78$), (b) conduct problems (e.g. has a temper, fights, lies: $\alpha_{\text{kindergarten}} = .65$, $\alpha_{\text{school}} = .50$), (c) hyperactivity/inattention (e.g. restless, distractible, uneven: $\alpha_{\text{kindergarten}} = .82$, $\alpha_{\text{school}} = .88$) as well as (d) peer problems (e.g. bullied, solitary, unpopular: $\alpha_{\text{kindergarten}} = .67$, $\alpha_{\text{school}} = .53$) which can be summed up to a total difficulty score ($\alpha_{\text{kindergarten}} = .83$, $\alpha_{\text{school}} = .80$). In contrast to the problem behaviors, (e) prosocial behavior has been added (e.g. helping, caring, sharing: $\alpha_{\text{kindergarten}} = .73$, $\alpha_{\text{school}} = .81$). Each of the subscales comprises five three-point likert items with the responses: “not true”, “somewhat true” and “certainly true”.

1.4.4. Attachment-Q-Sort

The AQS captures attachment relationships of children between 1 and 6 years in daily life situations. The 90 items characterize the relationship according to emotional security, stress reduction, support for exploration, affective sharing, positive attention, and physical contact. After a two-hour observation, the observer assigns the items to nine categories (each containing ten items) by using the Q-Sort method ranking from 9 (“very typical for the child”) to 1 (“very untypical for the child”) (forced assignment). The resulting AQS score is then computed by correlating the sorting of an individual observation with a sorting of an “ideal” secure relationship provided by experts.

1.4.5. Students' Reports on Feelings about School

The German instrument SRFS captures emotional and social experiences in school. Three subscales obtain students' reports on their (a) feeling of acceptance (13 items: $\alpha = .73$) measuring the extent to which a student feels emotionally supported and appreciated by the teachers, (b) attitudes toward school (14 items: $\alpha = .95$) evaluating the extent to which a student enjoys going to school and feels comfortable there, and (c) attitudes toward learning (13 items: $\alpha = .86$) assessing the extent to which a student is eager to

learn. Responses are recorded on a two-point scale (2 = “I agree” or 1 = “I don’t agree”). In addition, subscale (b) has been extended by two questions (response rates: “yes”, “no” or “sometimes”) on whether the teacher is perceived as “nice”, and whether the student “is pleased to see the teacher” (Harrison et al., 2007).

1.5. Data preparation, descriptive statistics and analysis strategy

1.5.1. Missing data

Of the total data set, 4% of the data were missing and in more than half of these cases, however, only one or two items per subject were missing. We thus used the *full information maximum likelihood* (FIML) for all further models allowing efficient model estimation based on all available data (Enders, 2010).

1.5.2. Descriptive statistics

According to Curran, West, and Finch (1996), item statistics (mean, standard deviation, skewness and kurtosis) for the kindergarten and elementary school sample were carried out (see Table 2). Only three items (items 23, 25, 26) appeared skewed with skewness values being greater than |2| and further item 26 with a kurtosis value greater than |7|. Univariate normality did not appear to be violated as a major pattern and was not taken into account for further statistical analyses, also because maximum likelihood statistics were proven to be quite robust to mild violations of normality (Jackson, Gillaspay, & Purc-Stephenson, 2009).

1.5.3. Data analysis strategy

All analyses were performed using the open-source software R 3.0.1 (R Development Core Team, 2013), using the package *lavaan* (Rosseel, 2012) to perform the CFA and measurement invariance tests, the package *semTools* (Pornprasertmanit, Müller, Schoemann, & Rosseel, 2013) for EFA model estimation with FIML and the package *nFactors* (Raïche & Magis, 2010) in order to extract the number of factors.

In order to conservatively confirm a measurement model (CFA) derived from the exploration of the factor structure (EFA), we split our sample into randomly selected sample-halves (Arlot & Celisse, 2010). To reduce bias due to sample selection, we randomly selected 100 training samples for the EFA (each $n = 435$) and 100 validation samples for the CFA (each $n = 436$), i.e., 100 EFA analyses were conducted on the training samples (one model per sample) and 100 CFA analysis on the validation samples. In the following, we report median factor loadings and medians of model fit statistics as well as the respective standard deviations within these 100 split-half sets in order to bootstrap the variability of the model estimates (Efron & Tibshirani, 1993). The loops necessary to run the models 100 times and to extract results were programmed with the R package *plyr* (Wickham, 2011).

To evaluate CFA models, we used several fit statistics such as root mean square error of approximation (RMSEA) including confidence interval (CI), standardized root mean square residual (SRMR), comparative fit index (CFI) and the Tucker–Lewis index (TLI) which tend to be unbiased by sample size (in contrast to χ^2 statistics, Bentler & Bonett, 1980; Jackson et al., 2009). In order to indicate acceptable model fits, we expected CFI and TLI values greater than .90 (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Hu & Bentler, 1999), RSMEA values up to .08 for acceptable fit (Hu & Bentler, 1999), as well as SRMR values less than .08. In addition, we explored R^2 values, factor loadings, and t -statistics. For model selection during the process of assessing measurement invariance, Bayesian information criterion (BIC) was applied of which the lowest BIC should indicate the preferred model (Bryk & Raudenbush, 1992). We also followed the recommendations of Cheung and Rensvold (2002), suggesting changes of the CFI less than or equal to $-.01$ ($\Delta CFI = CFI_{\text{constrained}} - CFI_{\text{unconstrained}}$) as indicators for acceptable

invariance model fit, which, in contrast to $\Delta\chi^2$, are not supposed to be affected by sample size. Therefore, chi squared statistics were not considered for model selection, but were reported.

2. Results

2.1. Examining the factor structure of the STRS applied in German-speaking samples

Based on Pianta’s (2001) three-factor structure of the STRS, we conducted a CFA for the whole sample based on a covariance-matrix. Factors but not residual variances were modeled to be correlated. For identification, factor loadings of all first indicator items were fixed to one. The model fit statistics for the present data were unacceptable: RMSEA = .08 (CI = .08–.09); TLI = .69; CFI = .72; SRMR = .10, $\chi^2(347) = 2417.32, p < .001$. In addition, low factor loadings and low values of R^2 indicated that subscales were weakly related to their factors. In order to improve the model, we subsequently excluded items with R^2 values $< .25$ as well as items with crossloadings, and ran further CFAs. However, also these CFAs failed to demonstrate satisfying model fits.

2.1.1. Exploration of factor structure (training samples)

Because the original STRS model did not fit the German data, we restarted the evaluations based on an exploratory approach on the test samples. In order to examine the number of factors, we applied parallel analyses (Horn, 1965) and optimal coordinates, a non-graphical solution to the scree test (Raïche, Walls, Magis, Riopel, & Blais, 2013). The parallel analyses and optimal coordinates revealed a three-factor solution in $n = 99$ samples and a two-factor solution in only one sample. In order to examine the factor loadings of the respective items of the three-factor solution, we applied oblique rotation (quartimax) because the factors appeared to be intercorrelated in several past studies using the STRS (Birch & Ladd, 1998; Gregoriadis and Tsigilis, 2008; Pianta & Stuhlman, 2004). Oblique rotated factor loadings and communality h^2 of each item and subscale are presented in Table 2. Following Floyd and Widaman’s (1995) suggestions, factor loadings greater than .40 (rounded) were considered to be meaningful (printed in bold; see Table 2), specifically because our sample size exceeded 300.

As a result of the EFA, Factor 1 was represented, in total, by seven substantial items; six of them (18, 20, 23, and 24–26) that had also described the scale conflict in the U.S. version of the STRS (called “original” in the following; Pianta, 2001). Furthermore, one item (15) which originally belonged to the closeness subscale was also allocated to Factor 1. In contrast, item 28 revealed crossloadings with Factor 1 and 2 and was thus dropped, according to Costello and Osborne (2005). Because the majority of items stemmed from the original STRS conflict scale, we interpreted Factor 1 as conflict scale. Factor 2 incorporated in total eight substantial items. Seven of them (1, 3–5, 7, 9, and 27) also belonged to the original STRS scale closeness; only one item (10) derived from the original STRS scale dependency. Thus, Factor 2 could be interpreted as measuring closeness. Factor 3 was not easy to describe on the basis of seven items in total. That is, three of the items (6, 8, 17) originated from the dependency subscale, and four items (11, 13, 16, 22) stemmed from the original conflict scale, having behavioral and/or communicative strategies in common to aggressively interrupt a relationship. As these behavioral patterns serve children’s autonomy in an inappropriate way and attachment theory considers them typical for dependent children (Sroufe, 2005; Sroufe et al., 1983) we labeled Factor 3 as dependency. In sum, the three factors of this German STRS suggest an adjusted three-factor model of Conflict (Factor 1), Closeness (Factor 2), and Dependency (Factor 3). Whereas, once again, most of the original items described factor

Table 2
Descriptive statistics and factor loadings of EFA and modified CFA model of the German version of the STRS.

Item: abbreviated item content	Overall sample statistics			Factor loadings for EFA				Factor loadings for CFA (modified model)		
	M (SD)	Skewness	Kurtosis	Conflict	Closeness	Dependency	<i>h</i> ²	Conflict St.Est. (<i>R</i> ²)	Closeness St.Est. (<i>R</i> ²)	Dependency St.Est. (<i>R</i> ²)
Conflict										
2: c. and I struggle with each other	1.61 (.95)	1.61	1.96	.20	−.28	.16	.14	–	–	–
11: c. easily becomes angry with me	1.70 (.89)	1.16	.64	.24	−.08	.52	.33	–	–	.66 (.44)
13: c. feels that I treat it unfairly	1.56 (.84)	1.59	2.25	.19	−.21	.48	.31	–	–	.74 (.55)
16: c. sees me as punishment/criticism	1.68 (.84)	1.20	1.12	.17	−.27	.44	.30	–	–	.69 (.48)
18: c. is angry after being disciplined	1.64 (.93)	1.53	1.64	.51	−.09	.07	.27	.59 (.35)	–	–
19: c. responds to me when it misbehaved	4.20 (.90)	−1.53	2.68	−.34	.24	.07	.18	–	–	–
20: c. drains my energy	1.64 (.96)	1.50	1.49	.70	.05	.06	.50	.67 (.45)	–	–
22: difficult day when c. has bad mood	1.67 (.90)	1.15	.48	.21	.03	.35	.17	–	–	–
23: c. feelings can be unpredictable	1.43 (.86)	2.18	4.26	.50	−.02	.23	.30	.61 (.37)	–	–
24: uncomfortable how c. and I get along	1.53 (1.00)	1.96	3.05	.46	−.22	.08	.27	.62 (.38)	–	–
25: c. cries to get something from me	1.36 (.78)	2.39	5.36	.72	.19	.01	.55	–	–	–
26: c. is manipulative with me	1.29 (.73)	2.80	7.56	.61	.04	.05	.38	.61 (.37)	–	–
Closeness										
1: affectionate, warm relationship with c.	4.44 (.78)	−1.62	3.02	−.26	.57	−.01	.39	–	–	–
3: upset c. will seek comfort from me	3.71 (1.12)	−.81	−.04	.01	.53	.01	.28	–	.59 (.35)	–
4: c. is uncomfortable with touch from me	1.44 (.80)	1.96	3.46	.07	− .52	.23	.33	–	–	–
5: c. values its relationship with me	4.32 (.77)	−1.25	2.21	−.04	.63	−.15	.42	–	.69 (.48)	–
7: when praising c. it beams with pride	4.54 (.70)	−1.73	3.45	.01	.54	.08	.30	–	–	–
9: c. shares information about itself	4.12 (1.03)	−1.19	.71	.24	.67	−.09	.51	–	–	–
12: c. tries to please me	2.52 (1.17)	.24	−.94	−.11	.19	.33	.16	–	–	–
15: easy to be in tune with c. feelings	4.02 (.93)	−.89	.34	− .54	.31	.05	.39	–	–	–
21: c. is copying my behavior	3.07 (1.01)	−.42	−.18	−.13	.29	.22	.15	–	–	–
27: c. shares its feelings/experiences	3.90 (1.05)	−1.02	.48	−.01	.65	.03	.42	–	.63 (.40)	–
28: interactions make me feel confident	4.18 (.89)	−1.10	1.08	−.49	.37	.01	.38	–	–	–
Dependency										
6: c. appears hurt when I correct it	2.46 (1.18)	.35	−1.07	−.05	−.07	.42	.18	–	–	–
8: c. reacts strongly to separation from me	2.52 (1.07)	.10	−.71	−.08	.30	.42	.27	–	–	–
10: c. is overly dependent on me	3.44 (.94)	−.45	.00	−.04	.64	.22	.46	–	.60 (.36)	–
14: c. asks needlessly for my help	1.89 (1.06)	1.14	.42	.30	.16	.24	.17	–	–	–
17: jealous c. when I concentrate on others	1.65 (.92)	1.36	1.18	.24	.14	.45	.28	–	–	–

Note. c.–child, item 4 and item 19 are reversed items, substantial EFA factor loadings rounded $\geq .40$ are printed in bold and included in the CFA.

1 and 2, factor 3 moves its semantic from the former description of a clingy relationship, to stress the tendency to escape from a clingy relationship by striving for autonomy. Although the item set of the German STRS scales are reduced and moved (see German items in Appendix), these scales are similar to Pianta's (2001) original STRS regarding their item loadings (see Table 2). Moreover, correlations of the factors were as follows: Conflict was negatively associated with closeness ($r = -.32$) and positively with dependency ($r = .40$); whereas dependency was negatively correlated with closeness ($r = -.10$).

2.1.2. Confirmation of the factor structure (validation samples)

To evaluate the factor structure derived by the EFAs, CFA analyses were carried out separately for the 100 random samples (validation samples). Once more, factors were allowed to correlate but not the residuals. The first model was unsatisfying, with the following averaged model fit indices: RMSEA = .08, ($SD = .003$), RMSEA CI = .07–.09, CFI = .81 ($SD = .01$), TLI = .79 ($SD = .01$) and SRMR = .09 ($SD = .004$), $\chi^2(206) = 760.40$, $p < .001$, even though it was improved by removing the following items stepwise: (a) items with R^2 lower than .20 (8, 6, 22, 17) and (b) items with substantial crossloadings with neighboring subscales (15, 1, 4, 25, 9, 7) regarding modification indices. Those items removed due to crossloadings were also conspicuous due to substantial residual correlations (items with residual correlations: 1 with 9, 20 and 27; 25 with 15 and 26; 9 with 7 and 27).

We conducted CFAs after each alternation of the model which improved RMSEA, CFI and TLI after each alternation finally resulting in adequate model fit statistics: RMSEA = .05 ($SD = .006$), RMSEA CI = .04–.07, CFI = .95 ($SD = .01$), TLI = .94 ($SD = .01$) and SRMR = .05 ($SD = .003$), $\chi^2(51) = 110.68$, $p < .001$. Thus, the final model for the German STRS comprised 12 items (see factor loadings in Table 2). The high factor loadings and large values for R^2 indicated that the selected items are highly related to its factors. Average factor correlations across the 100 samples (median) were as follows: Conflict was correlated about $r = -.40$ with closeness and about $r = .75$ with dependency and dependency and closeness were correlated about $r = -.46$. Further subscales provide satisfying reliabilities for the whole sample ($n = 871$) with an alpha of .75 equally for the conflict, closeness and dependency scale.

2.2. Invariance tests of the German STRS across context and gender

To evaluate whether group mean differences of gender and context can be meaningfully interpreted, several measurement invariance tests on the entire sample were conducted by sequentially adding model constraints in a series of hierarchically nested models (see Table 3).

2.2.1. Invariance across context

We first tested for configural invariance across context by applying the same factorial model across the three groups: kindergarteners, firstgraders, and secondgraders. Results revealed good model fit for the three groups of children (see Table 3). We then tested for weak factorial invariance by holding factor loadings equal across groups. Once again, model fit statistics were satisfying. Due to ΔCFI of .00 and decreased BIC, it can be concluded that the model showed weak factorial invariance across the three contexts (see Table 3).

Furthermore, to test for strong factorial invariance, intercepts were held equal across groups. A ΔCFI of $-.02$ revealed that intercepts were not equivalent across the three groups even though BIC was decreasing. An investigation of the intercepts revealed that item 3 and 5 of the closeness scale and item 23 of the conflict scale were highest in kindergarten and lowest in second grade. Thus,

their intercepts decreased steadily. Item 16 of the dependency scale showed an opposite pattern, whereby the intercept was lowest in kindergarten and highest in second grade. Thus, we tested partial strong invariance constraining intercepts to be equal across groups for all except the items 3, 5, 23 and 16 (loadings of these items were still constrained to be equal across groups). The model of partial strong invariance revealed a ΔCFI of $-.01$ and a decreased BIC compared to the weak factorial invariance model, as well as adequate model fit statistics (see Table 3).

When additionally holding residuals equal across groups, an increase in BIC and a ΔCFI of $-.03$ did not support the partial strict factorial invariance model. Thus, invariance test across contexts supported partial strong factorial invariance across kindergarteners, first graders and second graders.

2.2.2. Invariance across gender

Results revealed good model fit for the configural invariance model (see Table 3). After constraining the factor loadings, BIC decreased, $\Delta CFI = -.01$, revealing weak factorial invariance. Moreover, the model for strong factorial invariance was supported because BIC decreased and a $\Delta CFI = .00$. Strict factorial invariance was not supported, because $\Delta CFI = -.06$ and BIC increased and model fits were also not adequate. Thus, across boys and girls the STRS scales showed to be equivalent regarding factor loadings and intercepts (strong factorial variance).

2.3. The German STRS and its validation

In order to test for construct and criterion validity on the German STRS, separate structural equation models (SEM) for the kindergarten and school sample were fitted using subsamples of students for which the specific validity criteria were available. The latent scores of all three scales, conflict, closeness and dependency were regressed on one validation criteria per model, resulting in a model including the three latent variables of the STRS, one manifest variable (validation criteria) and three regression paths (see Table 4). Consequently, 24 different SEM models were run in total: one model (1) for the AQS, one model for each of the five subscales of the SRFS (1×5), two models (kindergarten and school sample) for each of the three subscales of the ICS, (2×3) and two models for each of the six subscales of the SDQ (2×6). All models yielded acceptable to good model fit indices in terms of RSMEA and CFI/TLI (see Table 4), so the validation models based on these measures has been proved to be successful, except for prosocial behavior in kindergarten.

2.3.1. Construct validity

In kindergarten, construct validity of the German STRS was confirmed by associations with AQS scores reflecting that the more secure the students had been evaluated in kindergarten, the closer ($\beta = .35$), the less conflictual ($\beta = -.33$) and less dependent ($\beta = -.25$) the teacher–student relationship appeared. In school, construct validity can be attested for closeness due to the positive association to students' reports on how they feel about their teachers at school ($\beta_s = .24$), revealing that close relationships were associated with good feelings. Negative coefficients of students' attitude toward the teacher ($\beta_s = -.31$ to $-.46$), however, were found in conflict relationships, revealing that students felt less accepted and had negative feelings toward their teachers. Most importantly, no significant association appeared for dependency, suggesting that students were not affected by their struggle in less autonomous relations (see Table 4).

2.3.2. Criterion validity

As expected, criterion validity is reflected by the fact that close relationships in kindergarten as well as in school were positively

Table 3
Model fit statistics of measurement invariance tests for the German STRS.

Model	χ^2	df	$\Delta\chi^2$	Δdf	RMSEA (CI)	CFI/TLI	ΔCFI	BIC
Model Context (g: kiga, 1st grade, 2nd grade)								
configural inv.	302.29 ^a	153	–	–	.06 (.05–.07)	.94/.93	–	25,393
weak factorial inv.	321.56 ^a	171	19.27	18	.06 (.05–.06)	.94/.94	.00	25,290
strong factorial inv.	404.60 ^a	189	83.04 ^a	18	.06 (.05–.07)	.92/.92	–.02	25,251
partial strong factorial inv.^a	366.67^a	181	45.11^b	10^b	.06 (.05–.07)	.93/.92	–.01^b	24,985
partial strict factorial inv. ^c	2886.28 ^a	198	2519.61 ^a	17	.07 (.06–.08)	.90/.90	–.03	25,017
Model Gender (g: boys, girls)								
configural inv.	206.12 ^a	102	–	–	.05 (.04–.06)	.96/.94	–	25,075
weak factorial inv.	221.61 ^a	111	15.49	9	.05 (.04–.06)	.95/.94	–.01	25,030
strong factorial inv.	243.70^a	120	22.09^a	9	.05 (.04–.06)	.95/.94	.00	24,991
strict factorial inv.	398.04 ^a	132	154.34 ^a	12	.07 (.06–.08)	.89/.89	–.06	25,064

Note. g = groups involved in the measurement invariance test, inv. = invariance, kiga = kindergarten; RMSEA = root mean square error of approximation, CI = confidence interval, CFI = comparative fit index, TLI = Tucker–Lewis index, AIC = Akaike information criterion, BIC = Bayesian information criterion, preferred model is printed in bold.

^a $p < .05$.

^a Intercepts of the items 3, 5, 23 and 6 were freely estimated across groups.

^b Compared to the weak invariance model.

associated with the extent of students' strength, such as social affiliation, prosocial behaviors, and popularity in the classroom ($\beta_s = .21$ to $.51$). In contrast, these strengths were negatively linked with conflict ($\beta_s = -.26$ to $-.52$) and dependency ($\beta_s = -.33$ to $-.62$; see Table 4).

However, children's difficulties, such as aggression, behavioral problems, or peer problems were negatively associated with closeness ($\beta_s = -.25$ to $-.55$) and positively with conflict ($\beta_s = .27$ to $.95$) and dependency ($\beta_s = .35$ to $.72$). Not surprisingly, criterion

validity of conflict further revealed that students in conflict relationships had bad attitudes toward school and learning ($\beta_s = -.30$ to $-.38$), whereas the feelings toward school and learning of children in close or dependent relationships were not affected. The general picture clearly reveals that close students are evaluated as pleasant and socially competent without behavioral difficulties, whereas students involved in conflictual or dependent relationships stand out due to their behavioral, emotional and social problems (for more details see Table 4).

Table 4
Validity criteria for the German STRS tested within several structural equation models: regression coefficients and model fit statistics.

Model	Regression coefficients			Model fit statistics		
	Closeness St.Est. (S.E.)/R ²	Conflict St.Est. (S.E.)/R ²	Dependency St.Est. (S.E.)/R ²	RMSEA (CI)	CFI/TLI	χ^2
Kindergarten models						
AQS ^a (external observation)	.35 ^{***} (.08)/.12	–.33 ^{***} (.08)/.11	–.25 ^{**} (.08)/.06	.06 (.03–.08)	.95/.93	92.41 [†]
ICS (teacher report)						
Aggression ^b	.08 (.09)/.01	.40 ^{***} (.08)/.16	.35 ^{***} (.07)/.12	.08 (.04–.08)	.94/.92	96.60 [†]
Popularity ^b	.21 [†] (.09)/.04	–.26 ^{**} (.08)/.07	–.15 [†] (.08)/.02	.05 (.03–.08)	.95/.93	90.22 [†]
Social Affiliation ^b	.45 ^{***} (.08)/.20	–.31 ^{***} (.08)/.09	–.35 ^{***} (.07)/.12	.06 (.04–.08)	.93/.91	102.03 [†]
SDQ (teacher report)						
Emotional Problems ^b	–.16 (.15)/.03	.21 (.13)/.05	.31 [†] (.13)/.09	.07 (.00–.11)	.94/.92	78.39
Behavioral Problems ^b	–.55 ^{***} (.13)/.30	.95 ^{***} (.03)/.91	.72 ^{***} (.08)/.52	.07 (.00–.11)	.95/.94	77.06
Hyperactivity ^b	–.19 (.15)/.04	.55 ^{***} (.10)/.31	.27 [†] (.13)/.07	.06 (.00–.11)	.95/.93	75.05
Peer Problems ^b	–.26 [†] (.14)/.07	.37 ^{**} (.12)/.14	.39 ^{**} (.12)/.15	.09 (.05–.13)	.90/.87	89.62 [†]
Prosocial Behavior ^b	.44 ^{**} (.15)/.19	–.52 ^{***} (.11)/.27	–.62 ^{***} (.10)/.38	.10 (.06–.14)	.88/.84	98.41 [†]
Total Difficulty ^b	–.42 ^{**} (.14)/.17	.78 ^{***} (.06)/.60	.59 ^{***} (.10)/.35	.06 (.00–.10)	.96/.94	73.60
School Models						
SRFS (student report)						
Feeling of Acceptance ^a	.08 (.12)/.01	–.46 ^{***} (.10)/.21	–.15 (.12)/.02	.08 (.04–.11)	.92/.90	91.38 [†]
Pleased to see the teacher ^a	.24 [†] (.12)/.06	–.31 ^{**} (.11)/.10	–.12 (.12)/.01	.07 (.03–.10)	.94/.92	83.50 [†]
Treated nicely by the teacher ^a	.24 [†] (.12)/.06	–.05 (.12)/.01	.05 (.12)/.01	.07 (.03–.10)	.93/.90	86.91 [†]
Attitude to School ^b	.10 (.12)/.01	–.38 ^{**} (.10)/.15	–.16 (.11)/.03	.07 (.02–.10)	.94/.92	82.65 [†]
Attitude to Learning ^b	.16 (.12)/.03	–.30 ^{**} (.11)/.09	–.07 (.12)/.01	.06 (.01–.09)	.94/.93	80.13 [†]
ICS (teacher report)						
Aggression ^b	–.25 ^{**} (.08)/.06	.60 ^{***} (.06)/.36	.43 ^{***} (.07)/.18	.07 (.05–.09)	.92/.90	117.24 [†]
Popularity ^b	.35 ^{***} (.08)/.12	–.42 ^{***} (.07)/.18	–.37 ^{***} (.07)/.14	.06 (.04–.08)	.94/.93	99.00 [†]
Social Affiliation ^b	.51 ^{***} (.07)/.26	–.41 ^{***} (.07)/.17	–.33 ^{***} (.08)/.11	.06 (.03–.08)	.95/.93	94.53 [†]
SDQ (teacher report)						
Emotional Problems ^b	–.07 (.12)/.01	.24 [†] (.11)/.06	.14 (.12)/.02	.07 (.03–.10)	.93/.90	87.26 [†]
Behavioral Problems ^b	–.18 (.12)/.03	.68 ^{***} (.07)/.47	.56 ^{***} (.08)/.32	.06 (.00–.09)	.95/.94	79.17 [†]
Hyperactivity ^b	–.23 [†] (.12)/.05	.62 ^{***} (.08)/.38	.29 ^{**} (.11)/.09	.08 (.05–.11)	.92/.89	93.28 [†]
Peer Problems ^b	–.34 ^{**} (.11)/.12	.27 [†] (.11)/.07	.36 ^{**} (.10)/.13	.06 (.02–.10)	.94/.92	82.22 [†]
Prosocial Behavior ^b	.42 ^{***} (.11)/.18	–.52 ^{***} (.09)/.27	–.47 ^{***} (.09)/.22	.06 (.00–.09)	.95/.94	77.52
Total Difficulty ^b	–.31 ^{**} (.12)/.09	.69 ^{***} (.07)/.48	.48 ^{***} (.09)/.23	.06 (.02–.10)	.95/.93	81.59 [†]

Note. St. Est, Standardized Estimates; S.E, Standard Error; R² refers to the explained variance of the latent factor due to the predictor variable, each row represents one SEM, degrees of freedom for all models = 60; RMSEA, root mean square error of approximation; CI, confidence interval; CFI, comparative fit index; TLI, Tucker–Lewis index.

^a Indicator for construct validity.

^b indicator for criterion validity.

[†] $p < .10$.

^{*} $p < .05$.

^{**} $p < .01$.

^{***} $p < .001$.

3. Discussion

The paper investigated the widely used STRS (Pianta, 2001) after its translation into German and its application in Germany and Austria. By running stepwise CFAs, we initially tested justification for the original three-factor model capturing closeness, conflict, and dependency. The deviation from the original model revealed unacceptable model fit indices and accordingly an exploration of the factorial structure of the items, from scratch, was necessary. We were able to provide an adjusted three-factor structure that is adequate for application of the STRS in Germany and Austria. This German STRS assures to validly assess student–teacher relationships on closeness, conflict, and dependency in those countries. Based on good model fits, the German STRS contains 12 of the original 28 items. Satisfying reliabilities of all three subscales outperformed even the weakness of the original dependency subscale that has been criticized due to its low alpha.

For the present German STRS, significant changes on the dependency subscale were important when the German translation had been tested, despite the notable item reduction. Whereas items from the original closeness and conflict subscales remained in the German STRS, items from the original dependency subscale were ineffective. Clearly, the original STRS captures features of a dependent relationship which adequately evaluated dependency also in Greece and the Netherlands (Gregoriadis & Tsigilis, 2008; Koomen et al., 2012), but not in the German speaking countries. In the German STRS, the dependency scale was filled with items from the original conflict subscale, which might more adequately describe behavioral consequences of a dependent relationship; i.e. the child tries to get released from it. This could be seen as the child's striving for autonomy resulting in conflicts with the teacher or vice versa. Thus, it is not surprising that the German dependency scale correlates closely with conflict.

The conceptual shift of dependency might be explained by different aspects of dependency being conspicuous in different individuals and cultures. From the individual point of view, Bowlby (1980) described dependency as a strong desire for close relationships because a dependent child “has no confidence that his attachment figure will be accessible” (Bowlby, 1980, p. 213). Consequently, the more dependent a child appears to be, the less close his relationship is. However, this negative dependency-closeness-association could only be confirmed for the presented applications of the STRS in German-speaking cultures, which are considered to treat people as highly individualistic and independent. Throughout independent cultures (Drugli & Hjemdal, 2013; Koomen et al., 2012; Webb & Neuharth-Pritchett, 2011), however, these associations were weakened by insufficient dependency scales. In contrast, in societies such as Greece, where interdependence is an important life and survival concept among its people, the dependency-closeness-correlations have clearly been positive.

Once more, from the individual point of view, Sroufe et al. (1983) suggest that the unaccomplished desire for closeness in dependent children might have “interfered with other developmental tasks, such as peer relations and environmental mastery” (p. 1626). Consequently, a child who cannot achieve his developmental tasks, such as establishing peer relations, might develop a certain amount of frustration. This might be the rationale as to why dependent children excessively show negative attention seeking strategies, hostile aggression, and behave impatiently and roughly toward their teachers in the classroom (Howes, Hamilton, & Matheson, 1994; Howes & Ritchie, 1999; Pianta & Nimetz, 1991; Sroufe, 2005; Sroufe et al., 1983). Thus, the more dependent a child is, the more she/he experiences conflict in relationships, even outside the home. This positive dependency-conflict-correlation could be found in the German STRS application as well as across cultures in which the STRS has been applied previously. From the cultural point of view,

independent cultures might stress the conflictual nature in dependent relations and therefore focus on coping strategies to overcome dependency, making the coping process conflictual, whereas interdependent cultures might stress the helpless nature of dependency and might focus on proximity to improve the relation. In other words, the present STRS reworking reveals dependency to be a culturally biased construct, whereas closeness and conflict might appear culturally invariant.

As a methodological approach, in order to justify construct and criterion validity, several SEMs were run to test validity criteria on the three subscales of the German STRS by using a multi-method approach: expert observations, teacher and student reports. Expert observations in kindergarten using the AQS clearly predicted all STRS subscales, confirming similar examinations carried out by Doumen et al. (2009), and Howes and Ritchie (1999) that showed STRS closeness to be positively associated and STRS conflict and dependency to be negatively related to AQS scores. These construct validations are promising because they indicate that a kindergarten teacher's report on his relationship quality is as good as an expert observation, which is a time-consuming appraisal of the student–teacher relationship in kindergarten. In the present study, student reports (in elementary school) about their feeling toward their teacher validated the STRS scales, conflict and closeness, as has been shown in past research (Doumen et al., 2009; Howes et al., 1998; Rey et al., 2007; Valeski & Stipek, 2001). Furthermore, several criteria for the German STRS subscales provided adequate evidence of criterion validity as the data sets provided rich links to many social-emotional domains. Thus, students' behavioral and emotional problems, aggression, peer conflicts, as well as popularity and prosocial behavior were strongly associated with the German STRS subscales, confirming prior research (Baker, 2006; Birch & Ladd, 1997; Birch & Ladd, 1998; Hamre & Pianta, 2001; Hughes & Cavell, 1999; Palermo et al., 2007).

Moreover, not only is the new German STRS economical regarding its application, it is also well suited to assess student–teacher relationships in diverse contexts such as across kindergarten and elementary school as well as with regard to gender. It enables researchers to meaningfully interpret differences across subgroups as well as over a period of time, which is important for longitudinal studies, specifically if the transition from kindergarten to elementary school is to be investigated. This study showed that the model, factor loadings, and intercepts were equal across gender and context, in specific ways. For example, applications of the German STRS confirmed strong factorial invariance of the STRS scales regarding gender as well as partially strong invariance across kindergarteners, first graders as well as second graders. Interestingly, the extent to which the teachers agreed to the various items describing the extent to which: *the child seeks comfort from the teacher when upset* (item 3), *the child values the relationship* (item 5) and *the child's behavior is unpredictable* (item 23) decreased steadily from kindergarten to second grade. From the developmental perspective, this pattern is well explicable, because children adopt more self-regulating strategies, therefore rely less on the teachers and focus more on other relationships such as peer relationships, and behavior patterns as well become more stable as they get older. On the contrary, the extent to which one agrees with the item 16 describing that *the child sees the teacher as punishment and criticism* steadily increases with time. This development might be due to the fact that when children get older, they pursue more and more autonomy, so they feel more restricted when the teacher gets involved in her/his actions. Thus, these four items have been shown to have a different meaning over time, also revealing that the students might be confronted with increasingly diverging challenges in elementary school, and thus develop changes of the relationship pattern toward their teachers.

The present German STRS was invariant for the kindergarten and school context and thus outperformed an earlier attempt by Mayr (2012), who proposed a German translation of the STRS within his teacher–child relationship measure for the kindergarten context only. The present work also improves on Mayr because appropriate statistical factor analysis, invariance tests and a robust statistical approach based on a bootstrapping procedure were applied. In sum, this research presents the first measurement model based on EFAs and CFAs, invariance tests and validity criteria for the German STRS providing good psychometric properties.

However, there are some limitations in this study. Construct validity was either assessed by expert observation or by student reports but not by both procedures at the same time which would make the validity approval even stronger. Moreover, it should be noted that the outcomes regarding dependency need to be further investigated in order to explore in more detail the extent to which dependency underlies cross-cultural variations. Because dependency was never related to a validity criterion stronger than conflict was, more specific validity criteria for dependency should be addressed in further studies in order to obtain a clearer picture on this issue. Thus, to generate a measure which is able to assess students' perception of dependency would be an enlightening future research topic. A further limitation of the present study was that subjects were not sampled randomly. Nevertheless due to balancing gender and SES the sample was comparable to the population regarding these characteristics. The sample cannot be seen as representative, because non-Caucasian children were underrepresented in the present study. Thus, results from the present study currently only should be generalized for Caucasian students. Though these are the overwhelming majority in Germany and Austria further research on culturally diverse children and teachers in these countries are needed.

Given that the student–teacher relationship predicts several developmental outcomes before and after school entry, the STRS may function as a screening instrument to discover relationship problems, which could help to prevent school adaptation problems at an early stage. From this perspective, the STRS is urgently required in the practical field of school psychology and counseling, which, over the past few decades, has shifted its focus from learning, to emotional processes, such as and student–teacher relationships.

Acknowledgments

This research was supported by grants from the Deutsche Forschungsgemeinschaft allocated to Lieselotte Ahnert (AH 55/13-1, AH 55/16-1) and Bettina Hannover (HA 2381/8-1, HA 2381/8-2) who conducted the project in Stendal, Cologne, Vienna and Berlin. Special thanks to Ilka Wolter, Sophie Muellerbauer and all other assistants and students who supported this research.

Appendix. German STRS items of the three subscales

Conflict [Konflikt]

- Item 18: Dieses Kind reagiert überhaupt nicht, wenn es ermahnt oder zurechtgewiesen wird.
 Item 20: Dieses Kind fordert meine ganze Kraft.
 Item 23: Die Launen dieses Kindes können mir gegenüber unberechenbar sein oder plötzlich wechseln.
 Item 24: Trotz meiner Bemühungen bin ich unzufrieden, wie ich mit diesem Kind zurechtkomme.
 Item 26: Dieses Kind ist ziemlich hinterhältig oder versucht mich zu manipulieren.

Closeness [Nähe]

- Item 3: Ist dieses Kind traurig, sucht es gern Trost bei mir.
 Item 5: Dieses Kind fühlt sich in meiner Gegenwart sehr wohl.

Appendix (Continued)

- Item 10: Dieses Kind hängt sehr an mir.
 Item 27: Dieses Kind teilt mir offen seine Gefühle und Gedanken mit.
 Dependency [Abhängigkeit]
 Item 11: Dieses Kind wird schnell ärgerlich auf mich.
 Item 13: Dieses Kind empfindet zumeist, ich würde es unfair behandeln.
 Item 16: Das Kind sieht mich als jemanden an, der vor allem maßregelt und kritisiert.

Note. Item number corresponds with the original STRS item number (Pianta, 2001), text in squared brackets refers to the corresponding German scale label.

References

- Ahnert, L., Harwardt-Heinecke, E., Kappler, G., Eckstein-Madry, T., & Milatz, A. (2012). Student–teacher relationships and classroom climate in first grade: How do they relate to students' stress regulation? *Attachment and Human Development, 14*(3), 249–263.
- Ahnert, L., Milatz, A., Kappler, G., Schneiderwind, J., & Fischer, R. (2013). The impact of teacher–child relationships on child cognitive performance as explored by a priming paradigm. *Developmental Psychology, 49*(3), 554–567. <http://dx.doi.org/10.1037/a0031283>
- Ahnert, L., Pinquart, M., & Lamb, M. E. (2006). Security of children's relationships with nonparental care providers: A meta-analysis. *Child Development, 74*(3), 664–679.
- Arbeau, K. A., Coplan, R. J., & Weeks, M. (2010). Shyness, teacher–child relationships, and socio-emotional adjustment in grade 1. *International Journal of Behavioral Development, 34*(3), 259–269.
- Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys, 4*, 40–79. <http://dx.doi.org/10.1214/09-ss054>
- Baker, J. A. (2006). Contributions of teacher–child relationships to positive school adjustment during elementary school. *Journal of School Psychology, 44*(3), 211–229.
- Behling, O., & Law, K. S. (2005). *Translating questionnaires and other research instruments: Problems and solutions* (3rd. ed.). Sage: Thousand Oaks, CA (print).
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588–606.
- 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.
- Birch, S. H., & Ladd, G. W. (1998). Children's interpersonal behaviors and the teacher–child relationship. *Developmental Psychology, 34*(5), 934–946.
- Booth, C. L., Kelly, J. F., Spieker, S. J., & Zuckerman, T. G. (2003). Toddlers' attachment security to child-care providers: The safe and secure scale. *Early Education and Development, 14*(1), 83–100.
- Bowlby, J. (1980). *Attachment and loss, vol. II: Separation, anxiety and anger*. London, UK: Hogarth.
- Bretherton, I. (1992). The origins of attachment theory: John Bowlby and Mary Ainsworth. *Developmental Psychology, 28*(5), 759–775.
- Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm, & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99–136). Washington, DC: American Psychological Association.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models in social and behavioral research: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Buyse, E., Verschueren, K., Verachtert, P., & van Damme, J. (2009). Predicting school adjustment in early elementary school: Impact of teacher–child relationship quality and relational classroom climate. *The Elementary School Journal, 110*(2), 119–141.
- Cairns, R. B., Leung, M.-C., Gest, S. D., & Cairns, B. D. (1995). A brief method for assessing social development: Structure, reliability, stability, and developmental validity of the interpersonal competence scale. *Behaviour Research and Therapy, 33*(6), 725–736.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255.
- Coplan, R. J., & Prakash, K. (2003). Spending time with teacher: Characteristics of preschoolers who frequently elicit versus initiate interactions with teachers. *Early Childhood Research Quarterly, 18*(1), 143–158.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation, 10*, 173–178.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to non-normality and specification error in confirmatory factor analysis. *Psychological Methods, 1*(1), 16–29.
- Davis, H. A. (2003). Conceptualizing the role and influence of student–teacher relationships on children's social and cognitive development. *Educational Psychologist, 38*(4), 207–234.
- 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.
- 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. <http://dx.doi.org/10.1080/00313831.2012.656697>
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.
- Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological assessment*, 7(3), 286–299.
- Fraire, M., Longobardi, C., & Sclavo, E. (2008). Contribution to validation of the student–teacher relationship scale (STRS Italian Version) in the Italian educational setting. *European Journal of Education and Psychology*, 1(3), 49–59.
- Glüer, M. (2012). *(Relationship quality and children's cooperation and educational engagement: A study in kindergarten and elementary school) Beziehungsqualität und kindliche Kooperations- und Bildungsbereitschaft: Eine Studie in Kindergarten und Grundschule*. Wiesbaden, Germany: VS.
- Greenfield, P. M., Keller, H., Fuligni, A., & Maynard, A. (2003). Cultural pathways through universal development. *Annual Review of Psychology*, 54(1), 461–490.
- Gregoriadis, A., & Tsigilis, N. (2008). Applicability of the student–teacher relationship scale (STRS) in the Greek educational setting. *Journal of Psychoeducational Assessment*, 26(2), 108–120.
- Hamilton, C. E., & Howes, C. (1992). A comparison of young children's relationships with mothers and teachers. *New Directions for Child and Adolescent Development*, 57, 41–59.
- 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.
- Harrison, L. J., Clarke, L., & Ungerer, J. A. (2007). Children's drawings provide a new perspective on teachers' child relationship quality and school adjustment. *Early Childhood Research Quarterly*, 22(1), 55–71.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185.
- 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.
- Howes, C., Hamilton, C. E., & Matheson, C. C. (1994). Children's relationships with peers: Differential associations with aspects of the teacher–child relationship. *Child Development*, 65(1), 253–263.
- Howes, C., Hamilton, C. E., & Philippsen, L. C. (1998). Stability and continuity of child–caregiver and child–peer relationships. *Child Development*, 69(2), 418–426.
- Howes, C., & Matheson, C. C. (1992). Contextual constraints on the concordance of mother–child and teacher–child relationships. *New Directions for Child and Adolescent Development*, 57, 25–40.
- Howes, C., & Ritchie, S. (1999). Attachment organizations in children with difficult life circumstances. *Development and Psychopathology*, 11(2), 251–268.
- 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.
- Hughes, J. N., & Cavell, T. A. (1999). Influence of the teacher–student relationship in childhood conduct problems: A prospective study. *Journal of Clinical Child Psychology*, 28(2), 173.
- Jackson, D. L., Gillaspay, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, 14(1), 6–23.
- Kennedy, J. H., & Kennedy, C. E. (2004). Attachment theory: Implications for school psychology. *Psychology in the Schools*, 41(2), 247–259.
- Klasen, H., Woerner, W., Wolke, D., Meyer, R., Overmeyer, S., Kaschnitz, W., et al. (2000). Comparing the German versions of the strengths and difficulties questionnaire (SDQ-Deu) and the child behavior checklist. *European Child and Adolescent Psychiatry*, 9(4), 271–276.
- 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.
- Ladd, G. W., & Price, J. M. (1987). Predicting children's social and school adjustment following the transition from preschool to kindergarten. *Child Development*, 58(5), 1168–1189.
- Mantzicopoulos, P., & Neuharth-Pritchett, S. (2003). Development and validation of a measure to assess head start children's appraisals of teacher support. *Journal of School Psychology*, 41(6), 431–451.
- Mashburn, A. J., & Pianta, R. C. (2006). Social relationships and school readiness. *Early Education and Development*, 17(1), 151–176.
- Mayr, T. (2012). Zur Erfassung von Erzieherin–Kind Beziehungen, Eine explorative Studie zum Einsatz der Student–Teacher Relationship Scale (STRS) in deutschsprachigen Kindertageseinrichtungen [Assessment of teacher–child relationships – Results of an explorative study of the student–teacher relationship scale (STRS) in German day care centers]. *Psychologie in Erziehung und Unterricht*, 59(2), 132–145.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543.
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY: Routledge.
- Murray, C., Murray, K. M., & Waas, G. A. (2008). Child and teacher reports of teacher–student relationships: Concordance of perspectives and associations with school adjustment in urban kindergarten classrooms. *Journal of Applied Developmental Psychology*, 29(1), 49–61.
- Niederhofer, H. (2000). *(Questionnaire to assess the parent–child relationship of children in the school age) Fragebogen zur Erfassung der Eltern–Kind–Beziehung im Schulalter*. (unpublished manuscript).
- O'Connor, E., & McCartney, K. (2006). Testing associations between young children's relationships with mothers and teachers. *Journal of Educational Psychology*, 98(1), 87–98.
- Palermo, F., Hanish, L. D., Martin, C. L., Fabes, R. A., & Reiser, M. (2007). Preschoolers' academic readiness: What role does the teacher–child relationship play? *Early Childhood Research Quarterly*, 22(4), 407–422.
- Pederson, D. R., & Moran, G. (1995). A categorical description of infant–mother relationships in the home and its relation to q-sort measures of infant–mother interaction. *Monographs of the Society for Research in Child Development*, 60, 111–132.
- Peisner-Feinberg, S. E., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., & Kagan, S. L. (2001). The relation of preschool child–care quality to children's cognitive and social development trajectories through second grade. *Child Development*, 72, 1534–1553.
- Peng, K., Nisbett, R. E., & Wong, N. Y. C. (1997). Validity problems comparing values across cultures and possible solutions. *Psychological Methods*, 2(4), 329–344.
- Pianta, R. C. (1992). Conceptual and methodological issues in research on relationships between children and nonparental adults. *New Directions for Child and Adolescent Development*, 57, 121–129.
- Pianta, R. C. (1994). Patterns of relationships between children and kindergarten teachers. *Journal of School Psychology*, 32(1), 15–31.
- Pianta, R. C. (1999). *Enhancing relationships between children and teachers* (2nd ed.). Washington, DC: American Psychological Association.
- Pianta, R. C. (2001). *STRS. Student–teacher relationship scale*. Lutz, FL: Psychological Assessment Resources, Inc.
- Pianta, R. C., Hamre, B. K., & Stuhlman, M. (2003). Relationships between teachers and children. In W. Reynolds, & G. Miller (Eds.), *Handbook of psychology (vol. 7) Educational psychology* (pp. 199–234). Hoboken, NJ: Wiley.
- Pianta, R. C., & Nimetz, S. L. (1991). Relationships between children and teachers: Associations with classroom and home behavior. *Journal of Applied Developmental Psychology*, 12(3), 379–393.
- Pianta, R. C., Nimetz, S. L., & Bennett, E. (1997). Mother–child relationships, teacher–child relationships, and school outcomes in preschool and kindergarten. *Early Childhood Research Quarterly*, 12(3), 263–280.
- 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.
- Pianta, R. C., Steinberg, M. S., & Rollins, K. B. (1995). The first two years of school: Teacher–child relationships and deflections in children's classroom adjustment. *Development and Psychopathology*, 7(2), 295–312.
- Pianta, R. C., & Stuhlman, M. W. (2004). Teacher–child relationships and children's success in the first years of school. *School Psychology Review*, 33(3), 444–458.
- Pornprasertmanit, S., Miller, P., Schoemann, A., & Rosseel, Y. (2013). *semTools: Useful tools for structural equation modeling (R package Version 0.4-0) [Computer Software]*. Retrieved from <http://cran.r-project.org/web/packages/semTools/>
- Posada, G., Gao, Y., Wu, F., Posada, R., Tascon, M., Schöelmerich, A., et al. (1995). The secure-base phenomenon across cultures: Children's behavior, mother's preferences, and experts' concepts. *Monographs of the Society for Research in Child Development*, 60(2/3), 27–48.
- Raïche, G., & Magis, D. (2010). *nFactors: Parallel analysis and non graphical solutions to the Cattell's scree test (R package Version 2.3.1.) [Computer Software]*. Retrieved from <http://cran.r-project.org/web/packages/nFactors/index.html>
- Raïche, G., Walls, T. A., Magis, D., Riopel, M., & Blais, J.-G. (2013). Non-graphical solutions for Cattell's scree test. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 9(1), 23–29. <http://dx.doi.org/10.1027/1614-2241/a000051>
- Rauer, W., & Schuck, K. D. (2004). *(Students' report on feelings about school) Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen*. Göttingen, Germany: Beltz.
- R Development Core Team. (2013). *R: A language and environment for statistical computing (Version 3.0.1) [Computer Software]*. Vienna, Austria: R Foundation for Statistical Computing.
- 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.
- Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, 81(4), 493–529.
- Rosseel, Y. (2012). *lavaan: An R package for structural equation modeling. Journal of Statistical Software*, 48(2), 1–36.
- Rudasill, K. M., & Rimm-Kaufman, S. E. (2009). Teacher–child relationship quality: The roles of child temperament and teacher–child interactions. *Early Childhood Research Quarterly*, 24(2), 107–120.
- Saal, F., Downey, R., & Lahey, M. (1980). Rating the ratings: Assessing the psychometric quality of rating data. *Psychological Bulletin*, 88(2), 413–428.
- Spilt, J. L., Hughes, J. N., Wu, J. Y., & Kwok, O.-M. (2012). Dynamics of teacher–student relationships: Stability and change across elementary school and the influence on children's academic success. *Child Development*, 83(4), 1180–1195.

- Sroufe, L. A. (2005). Attachment and development: A prospective, longitudinal study from birth to adulthood. *Attachment and Human Development*, 7(4), 349–367.
- Sroufe, L. A., Fox, N. E., & Pancake, V. R. (1983). Attachment and dependency in developmental perspective. *Child Development*, 54(6), 1615–1627.
- Statistical Federal Office Destatis. (2011). *Sample survey of income and expenditure*. Retrieved from https://www.destatis.de/EN/FactsFigures/SocietyState/IncomeConsumptionLivingConditions/IncomeReceiptsExpenditure/Tables/IncomeExpenditure_NetIncome.html
- Statistical Federal Office Destatis. (2012). *Persons with migrant a background*. Retrieved from <https://www.destatis.de/EN/FactsFigures/SocietyState/Population/MigrationIntegration/PersonsMigrationBackground/Tables/MigrantStatusSex.html>
- Statistik Austria. (2012). (Tables EU-SILC 2011. Income, poverty and living circumstances) Tabellenband EU-SILC 2011. Einkommen, Armut und Lebensbedingungen. http://www.statistik.at/web_de/statistiken/soziales/haushalts-einkommen/index.html
- Statistik Austria. (2013). (Table persons with migration background since 2008) Bevölkerung mit Migrationshintergrund seit 2008. Retrieved from http://www.statistik.at/web_de/statistiken/bevoelkerung/bevoelkerungsstruktur/bevoelkerung_nach_migrationshintergrund/069443.html
- 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.
- Valeski, T. N., & Stipek, D. J. (2001). Young children's feelings about school. *Child Development*, 72(4), 1198–1213.
- van IJzendoorn, M. H., & Sagi, A. (2008). Cross-cultural patterns of attachment: Universal and contextual dimensions. In J. Cassidy, & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (pp. 880–905). New York, NY: Guilford.
- Waters, E. (1995). The attachment Q-Set (Version 3.0). *Monographs of the Society for Research in Child Development*, 60, 234–246.
- Webb, M. L., & Neuharth-Pritchett, S. (2011). Examining factorial validity and measurement invariance of the student–teacher relationship scale. *Early Childhood Research Quarterly*, 26(2), 205–215.
- Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, 40(1), 1–29.
- Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, 28(3), 263–311.
- Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). Washington, DC: American Psychological Association.
- Woerner, W., Becker, A., & Rothenberger, A. (2004). Normative data and scale properties of the German parent SDQ. *European Child and Adolescent Psychiatry*, 13(2), ii/3–ii/10.