UvA-DARE (Digital Academic Repository)

Understanding cross-cultural differences in affective teacher student relationships: A comparison between Dutch and Chinese primary school teachers and students

Chen, M.; Zee, M.; Koomen, H.M.Y.; Roorda, D.L.

DOI

[10.1016/j.jsp.2019.07.011](https://doi.org/10.1016/j.jsp.2019.07.011)

Publication date

2019

Document Version

Final published version

Published in

Journal of School Psychology

[Link to publication](#)

Citation for published version (APA):

Chen, M., Zee, M., Koomen, H. M. Y., & Roorda, D. L. (2019). Understanding cross-cultural differences in affective teacher student relationships: A comparison between Dutch and Chinese primary school teachers and students. *Journal of School Psychology*, 76, 89-106. <https://doi.org/10.1016/j.jsp.2019.07.011>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the Library of the University of Amsterdam (<https://dare.uva.nl>)

Understanding cross-cultural differences in affective teacher-student relationships: A comparison between Dutch and Chinese primary school teachers and students

Mengdi Chen*, Marjolein Zee, Helma M.Y. Koomen, Debora L. Roorda

Research Institute of Child Development and Education, University of Amsterdam, P.O. Box 15776, NL-1001 NG Amsterdam, the Netherlands

ARTICLE INFO

Action Editor: Nate von der Embse

Keywords:

Cross-cultural comparisons
Individualism and collectivism
Student and teacher perceptions
Teacher-student relationships
Upper elementary school

ABSTRACT

The affective quality of teacher-student relationships (TSRs) has mostly been studied in Western contexts and little is known how findings generalize to Eastern contexts. Therefore, the present study examined measurement invariance, mean differences, and agreement in teachers' and students' perceptions of relationship quality between the Netherlands and China. The sample consisted of 789 primary school students (51% girls) and 35 teachers (80% female) from the Netherlands, and 587 primary school students (53% girls) and 14 teachers (93% female) from Zhejiang, China. Both teachers and students reported about the quality of their mutual relationships. Structural equation modeling showed that teachers' perceptions of closeness, conflict and dependency, as well as students' perceptions of closeness and conflict in the relationship, reached partially strong invariance. Chinese students perceived more closeness and less conflict with teachers than Dutch students. Chinese teachers also reported less conflict but comparable levels of closeness and dependency as Dutch teachers. Chinese teacher-student dyads had higher agreement on closeness but lower agreement on conflict than Dutch teacher-student dyads. Implications and suggestions for future research are discussed.

1. Introduction

The affective quality of teacher-student relationships (TSRs) has frequently been found to be important for children's social-emotional and academic development (for meta-analytic overviews, see [Lei, Cui, & Chiu, 2016](#); [Lei, Cui, & Chiu, 2018](#); [Roorda, Jak, Zee, Oort, & Koomen, 2017](#); [Roorda, Koomen, Spilt, & Oort, 2011](#)). Despite the importance of TSRs for children's school functioning, these relationships have been primarily studied in Western contexts and less often in Eastern contexts ([Bear et al., 2014](#); [Jia et al., 2009](#)). Still, cultural contexts may have an important influence on the development of TSRs ([Pianta, Hamre, & Stuhlman, 2003](#)). Findings of studies in Western, individualistic countries, which value the independence of individuals ([Markus & Kitayama, 1991](#)), may not necessarily fit to Eastern, collectivistic countries, which stress the fundamental connectedness to each other ([Markus & Kitayama, 1991](#); [Oyserman, Coon, & Kemmelmeier, 2002](#)). To get a better view of possible cross-cultural differences in the quality of TSRs, this study aimed to compare TSRs across the Netherlands (an individualistic country) and China (a collectivistic country).

Furthermore, previous studies in Western contexts usually found only low to moderate agreement between teachers' and students' perceptions of their mutual relationships ([Hughes, 2011](#); [Hughes & Cavell, 1999](#); [Rey, Smith, Yoon, Somers, & Barnett, 2007](#)). We do

* Corresponding author.

E-mail addresses: M.Chen@uva.nl (M. Chen), M.Zee@uva.nl (M. Zee), H.M.Y.Koomen@uva.nl (H.M.Y. Koomen), D.L.Roorda@uva.nl (D.L. Roorda).

<https://doi.org/10.1016/j.jsp.2019.07.011>

Received 29 October 2018; Received in revised form 12 June 2019; Accepted 23 July 2019

Available online 09 August 2019

0022-4405/ © 2019 Society for the Study of School Psychology. Published by Elsevier Ltd. All rights reserved.

not yet know whether this finding also generalizes to Eastern contexts. Therefore, the present study included both teachers' and students' perceptions of their relationships and also investigated informant agreement both in the Netherlands and China.

1.1. Cultural values and its impact on interpersonal relationships

The impact of culture on children's development is acknowledged by different theories, for example, the developmental systems model of Pianta et al. (2003). According to this model, TSRs are embedded in developmental systems and affected by external influences such as cultural values and norms. Cultural values and norms, such as general beliefs about the development of children, shape teachers' and students' expectations about each other and, hence, affect the way students and teachers relate to one another (Pianta et al., 2003).

With regard to cultural values, a prevailing practice is to make the distinction between individualistic and collectivistic cultures (Hofstede, 1980; Oyserman et al., 2002). According to the Individualism-Collectivism theory, people in an individualistic culture tend to be independent from their in-groups (e.g. family, tribe, and nation) and guide their behaviors on the basis of their own goals (Triandis, 2001; Triandis, 2018; Triandis, Bontempo, Villareal, & Asai, 1988). On the contrary, people in a collectivistic culture tend to be dependent on their in-groups and guide their behaviors on the basis of in-group norms (Triandis, 2001; Triandis et al., 1988).

The difference in individualism and collectivism is theorized to be associated with core psychological constructs such as relationships with others (Oyserman & Lee, 2008; Triandis, 2018). More specifically, interpersonal interdependence is assumed to be higher in collectivistic cultures than in individualistic cultures, where there is more distance and detachment (Triandis, 2018). Furthermore, in terms of preferred conflict-solving strategies, individuals in collectivistic cultures tend to avoid disturbance and confrontation by using compromise and disengagement, whereas people in individualistic cultures tend to adopt direct and assertive strategies (Chen & French, 2008). Empirical studies also found evidence supporting differences between individualistic and collectivistic cultures with regard to peoples' cognition and relationships with others (for meta-analytic overviews, see Oyserman & Lee, 2008), also in the field of education (Imada, 2012). Therefore, in the present study, we compared the affective quality of TSRs between individualistic and collectivistic cultures.

1.2. Cultural difference from the perspective of attachment theory

Attachment theory (Bowlby, 1980) helps to further understand affective relationships in cross-cultural contexts. This theory was initially developed to explain how parent-child relationships serve as a context for children's development. According to attachment theory (Bowlby, 1980), children would regard their mothers as a safe haven and seek proximity with their mothers when feel threatened or hurt. Mothers' sensitive responses (i.e., perceiving children's needs accurately and responding appropriately and timely towards them) are considered to form a secure base for children to explore the environment (Ainsworth, Blehar, Waters, & Wall, 1978). As a result, children who are securely attached become more socially and emotionally competent than children who are insecurely attached (Ainsworth et al., 1978).

Attachment relationships, however, may be defined or understood differently in different cultural contexts (Beyazkurk & Kesner, 2005; Keller, 2013; Mizuta, Zahn-Waxler, Cole, & Hiruma, 1996; Oyserman et al., 2002). In a systematic review, Rothbaum, Weisz, Pott, Miyake, and Morelli (2000) suggested that Japanese mothers interpreted sensitivity as being responsive to children's social engagement and as a way to stimulate interdependence and social harmony. In contrast, American mothers interpreted sensitivity as being responsive to children's need for independence, which would stimulate children's autonomous exploration and increase their autonomy. Therefore, the development of attachment relationships is likely to differ across Western and Eastern cultures.

Since the early 1990s, attachment theory has been argued to apply to teacher-student relationships as well (Pianta, 1992; Verschuere, 2015; Verschuere & Koomen, 2012). Unlike the attachment bond between children and parents, relationships between teachers and students are usually not exclusive, nor durable, and therefore may not be considered as a "full-fledged" attachment bond (Verschuere & Koomen, 2012). However, a teacher may serve as a temporary attachment figure and can also function as a secure base and safe haven for children, which enables students to comfortably explore the learning environment (Pianta et al., 2003; Verschuere & Koomen, 2012). In contrast, negative TSRs may harm students' ability to use teachers as a safe haven and secure base, which interfere with their attempts to explore the school environment (Pianta et al., 2003; Verschuere & Koomen, 2012). As TSRs seem to share important similarities with child-family relationships, findings with regard to cultural differences in child-family relationships may help us to gain insight about cultural differences in TSRs (Pianta et al., 2003).

1.3. Cross-cultural differences in family-child relationships

In line with the developmental systems model (Pianta et al., 2003) and attachment theory (Bowlby, 1980), some empirical studies have found evidences for cross-cultural differences in affective family-child relationships. For instance, Dwairy and Achoui (2010) found that family-child connectedness was lower in Western countries than in Eastern countries. In addition, American mothers were found to maintain less intimacy with their children than Indonesian mothers (French, Rianasari, Pidada, Nelwan, & Buhrmester, 2001) and Japanese mothers (Dennis, Cole, Zahn-Waxler, & Mizuta, 2002). With regard to the degree of conflict in interpersonal relationships, youth in the United States reported more conflict with family members than Indonesian youth (French et al., 2001). Lastly, Dennis et al. (2002) found that American mother-preschooler dyads emphasized more individuality in their conversations, whereas Japanese dyads emphasized more mutuality. Taken together, these empirical studies seem to imply that there are substantial differences in family-child relationships between individualistic and collectivistic cultures. Considering the assumed similarities

between family-child and teacher-student relationships, these findings may indicate that there are cultural differences in TSRs as well.

1.4. Cross-cultural differences in teacher-student relationships

As far as we know, there are only five published studies that actually investigated cross-cultural differences in affective teacher-student relationships. Studies based on attachment theory often distinguish between three dimensions of affective relationship quality, namely closeness (i.e., the degree of warmth and openness in the relationship), conflict (i.e., the discordance in the relationship) and dependency (i.e., a child's overreliance and possessiveness in the relationship; Pianta, 2001; Verschueren & Koomen, 2012). In terms of the degree of closeness and teacher support in the relationship, previous studies found higher relationship quality in collectivistic cultures from both teachers' and students' perceptions (Beyazkurk & Kesner, 2005; Jia et al., 2009; Yang et al., 2013). For instance, Yang et al. (2013) investigated third to twelfth grade students in China and the United States and found that Chinese students perceived higher closeness with teachers compared to American students. In addition, Beyazkurk and Kesner (2005) found that Turkish teachers in kindergarten and primary schools reported more closeness with their students than elementary school teachers in the United States. Jia et al. (2009) also found that Chinese students in urban middle schools perceived higher levels of teacher support compared to their counterparts in America. However, contradictory results were also found. That is, Aldridge and Fraser (2000) found no significant differences in the degree of student-reported teacher support between junior high school students (13–15 years old) from Australia and from Taiwan. Bear et al. (2014) even found that American elementary school teachers reported higher closeness than their Chinese counterparts, whereas students reported equally favorable relationships with teachers across countries.

In terms of the degree of conflict in TSRs, Beyazkurk and Kesner (2005) found no significant differences in teachers' perceptions of conflict between Turkey and the United States. With regards to dependency, as far as we know, no study has yet examined cross-cultural differences in this dimension. However, several studies in Greece seem to imply that dependency may have positive connotations in collectivistic cultures, whereas in individualistic cultures dependency mostly is considered as a negative relationship dimension (Gregoriadis & Grammatikopoulos, 2014; Gregoriadis & Tsigilis, 2008; Tsigilis, Gregoriadis, & Grammatikopoulos, 2018; Tsigilis, Gregoriadis, Grammatikopoulos, & Zachopoulou, 2018).

Altogether, previous studies found preliminary evidence that there might be variations in the affective quality of TSRs between individualistic and collectivistic cultures. Considering the limited number of empirical studies available, however, more research is needed before we can make inferences about cultural differences in TSRs. Moreover, as the world is becoming more and more closely related, it seems particularly important to establish mutual understandings and get insights from different cultural values. Therefore, the present study aimed to investigate the cultural differences in affective TSRs.

1.5. Agreement in teachers' and students' relationship perceptions

According to attachment theory, one of the key components of affective TSRs is teachers' and children's mental representation of the relationship (Pianta et al., 2003). Mental representation refers to each relationship partner's (i.e., the teacher or the student) feelings, emotions, and thoughts about him or herself, the other, and the mutual relationship (Pianta et al., 2003). As mental representations are formed based on attachment histories with parents, previous teachers or students, teachers and students may develop different mental representations of their mutual relationship. Therefore, their relationship perceptions may not always agree with each other (Pianta et al., 2003). In line with this assumption, empirical studies have found only low to moderate agreement in teachers' and students' relationship perceptions, with associations ranging from 0.21 to 0.59 (Hughes, 2011; Jellesma, Zee, & Koomen, 2015; Koomen & Jellesma, 2015; Rey et al., 2007), and some associations being even non-significant (Hughes, 2011; Hughes & Cavell, 1999). However, these findings were largely found in Western contexts and we do not know whether they also apply to Eastern contexts. Therefore, we also examined whether there are cross-cultural differences in the agreement between informants.

1.6. Cultural differences between China and the Netherlands

The present study aimed to compare the quality of TSRs between the Netherlands, an individualistic country, and China, a collectivistic country. These two countries are of particular interest because of their striking differences on three relationship-related dimensions of Hofstede's cultural dimensions theory (Hofstede, Hofstede, & Minkov, 2010). These dimensions are individualism (i.e., the degree of interdependence among members of a society), power distance (i.e., the extent that members accept the unequally distributed power in the society) and masculinity (i.e., the degree a society is driven by competition and achievement). More specifically, Dutch society highly values the independence of individuals (scores 80 on individualism), expects low power distance between people (scores 38 on power distance), and pays more attention to life-work balance instead of competition and achievement (scores 14 on masculinity; Hofstede et al., 2010). On the contrary, Chinese society values the interdependence among members (scores 20 on individualism), generally accepts subordinate-superior relationships among people (scores 80 on power distance), and cares much about competition and achievement (scores 66 on masculinity; Hofstede et al., 2010). As the degree of individualism, power distance and masculinity in a society can shape the way that teachers and students interact with each other (Hofstede et al., 2010), there may be cultural differences in TSRs between countries who have large differences in these dimensions. Therefore, the strikingly cultural differences between Dutch and Chinese society make them good candidates to make cross-cultural comparison on TSRs.

The few existing cross-cultural studies on TSRs mostly focused on comparisons between the United States and China (Bear et al., 2014; Jia et al., 2009; Yang et al., 2013). Following the literature, the present study also included China as a representative of collectivistic cultures, as China is one of the largest Asian countries with a large population. However, instead of the U.S., we focused on the Netherlands as a representative of individualistic cultures. The Netherlands had similar scores as the U. S. on individualism (80 in the Netherlands versus 91 in the U.S.) and power distance (38 in the Netherlands versus 40 in the U.S.). With regard to the scores on masculinity, however, the discrepancy is much larger between China (66) and the Netherlands (14), than the discrepancy between China and the United States (62). Therefore, the comparison between China and the Netherlands may be even more interesting than the comparison between China and the U.S.

1.7. The present study

The present study aimed to compare Dutch and Chinese third to sixth graders' affective relationships with their teachers, and included both students' and teachers' perceptions of relationship quality. We used the Student Teacher Relationship Scale (STRS, Pianta, 2001) and Student Perception of Affective Relationship with Teacher Scale (SPARTS; Koomen & Jellesma, 2015) to measure teachers' and students' relationship perceptions. These instruments have not been used in a cross-cultural comparison between the Netherlands and China before. As a meaningful cross-cultural comparison is only possible when adequate measurement invariance is established (Chen, 2008; Steenkamp & Baumgartner, 1998), we first checked for measurement invariance of the instruments across the Netherlands and China. This is a necessary step because in the absence of measurement invariance, systematic group differences in teachers' and students' relationship perceptions are not completely attributable to valid cultural differences, but also to measurement artifact (Jak, Oort, & Dolan, 2014). In other words, conducting cross-cultural comparisons without testing measurement invariance increases the risk of comparing chopsticks with forks (Chen, 2008). Therefore, we first tested measurement invariance of the STRS and SPARTS across the two countries. Second, we investigated cultural differences between the Netherlands and China in both teachers' and students' relationship perceptions. Based on previous studies about TSRs (Beyazkurk & Kesner, 2005; Jia et al., 2009; Yang et al., 2013), we hypothesized that Chinese students and teachers would report more closeness than their Dutch counterparts. Based on previous studies on family-child relationships (Dennis et al., 2002; French et al., 2001), we also expected to find less conflict and more dependency in the relationships between Chinese teachers and students compared to their Dutch counterparts. Third, we examined whether there were cross-cultural differences in the degree of agreement in relationship perceptions between teacher-student dyads from the Netherlands and China. As cultural differences in informant agreement about TSRs are hardly examined in previous research, we were not able to draw sound hypotheses and therefore kept this research question exploratory.

2. Method

2.1. Participants

2.1.1. Dutch sample

The Dutch student sample consisted of 789 students (51% girls) from eleven elementary schools in the Netherlands. In total, 43 classes participated in this study. On average, there were 23 students ($SD = 4.22$) in each class, ranging from 16 to 29 students per class, except for one uncharacteristically small class with only eight students. The Dutch students were in third ($n = 170$), fourth ($n = 218$), fifth ($n = 199$) and sixth ($n = 202$) grade, respectively. They had a mean age of 10.49 years ($SD = 1.23$), ranging from eight to 14 years old. Most of the students identified themselves (72%) and their parents (66% mothers and 64% fathers) as belonging to the Dutch ethnic group. Other students identified themselves and their parents as belonging to a minority group (i.e., Moroccan, Surinamese, Turkish, and other ethnic groups).

The Dutch teacher sample included 35 teachers (80% female), who reported about their relationship quality with eight randomly selected students from their classes. We decided to select eight students per class because it would be overly burdensome for teachers to fill in questionnaires for more students, and would thus compromise their willingness to participate. Moreover, as suggested by Snijders and Bosker (1999), high intra-class correlations may make it less beneficial to include whole classes.¹ Accordingly, teacher reports were available for 269 students (51% girl students). Due to time limits or illness, eight teachers reported about their relationships with less students than planned (ranging from five to seven students per teacher), resulting in missing teacher reports for eleven of the selected students. The teachers had on average 14.01 years of teaching experience ($SD = 11.68$; ranging from one to 43 years). Eight teachers worked full time (i.e., > 32 h per week), nine teachers worked four days per week, and 18 teachers worked three days per week.

¹ We did not find significant mean differences between the selected students and the total student sample on age ($t(787) = 1.04, p = .297$ for Dutch sample, and $t(567) = -0.28, p = .777$ for Chinese sample) and ethnicity ($t(787) = 0.61, p = .542$ for Dutch sample and $t(585) = 0.13, p = .896$ for Chinese sample). Furthermore, we compared the observed mean scores of students' relationships perceptions between the selected students and the total student sample. No significant mean differences between the selected students and the total sample were found for student-reported closeness ($t(787) = -0.65, p = .516$ for Dutch sample and $t(586) = -0.95, p = .341$ for Chinese sample) and conflict ($t(787) = 0.77, p = .438$ for Dutch sample and $t(586) = 1.14, p = .253$ for Chinese sample). Given these results, the selected students can be considered representative of the total sample.

2.1.2. Chinese sample

The Chinese student sample consisted of 587 students (53% girls) from three elementary schools in Zhejiang, China. In total, 14 classes participated in this study. On average, there were 42 students in each class ($SD = 5.25$), ranging from 34 to 52 students per class. The Chinese students were also in third ($n = 127$), fourth ($n = 220$), fifth ($n = 119$) and sixth ($n = 122$) grade. They had a mean age of 11.49 years old ($SD = 1.29$), ranging from nine to 14 years old. Chinese students' ethnicity was divided into two categories, namely Han (i.e., the major ethnic group in China) and minority (including 55 minor ethnic groups in China, such as She Zu). Most of the students identified themselves (94%) and their parents (96% fathers and 95% mothers) as belonging to the Han group.

The Chinese teacher sample included 14 teachers (93% females), who also reported about their relationship quality with eight randomly selected students from their classes. The Chinese teachers all completed the entire questionnaire and, hence, teacher reports were available for 112 students (50% girl students). The Chinese teachers had on average 9.65 years of teaching experience ($SD = 6.60$; ranging from two to 23 years). Twelve teachers worked full time, one teacher worked four days per week, and one teacher worked two days per week.

2.2. Procedure

2.2.1. Dutch sample

We received approval from the ethical committee of the University of Amsterdam (file number: 2016-CDE-7243) to carry out the data collection in the Netherlands. Schools were contacted via telephone and email. In case schools agreed to participate, teachers received information letters explaining the aim of this study as well as an informed consent form. The parents of the students also received an information letter and could object to their children's participation in this study.

Data were collected by trained assistants. During school visits in February and March 2017, all students in the classroom were asked to fill out a questionnaire about background variables, their relationships with their teachers, and some other topics that were beyond the scope of the present study. In case students had multiple part-time teachers, they were instructed to report on their relationships with the teacher teaching on most school days. The trained assistants explained the procedure and answered students' questions. Teachers were asked to leave the classroom while students were completing the questionnaires to encourage students' honest responses. It took the students around 30 min to finish the whole questionnaire. The total response rate was 99% and non-participation was mainly due to students' absence.

Around the same time of the school visits, teachers were emailed and asked to complete a digital questionnaire, measuring their relationship quality with eight randomly selected students (four boys and four girls). It took the teachers around 40 min to complete the whole questionnaire and they were asked to finish the questionnaire within two weeks.

2.2.2. Chinese sample

Because China has no official ethical committee, an independent senior researcher in China was invited to review the proposal and she agreed that the data collection procedure complied with the laws and rules in China. Principals in targeted schools were contacted via telephone to ask whether they were willing to let their schools participate. Similar to the Dutch procedure, teachers and parents received information letters and were asked to confirm their participation (teachers) or to let know if they objected against their children's participation (parents). Different from Dutch primary school teachers, who teach all subjects, Chinese primary school teachers often teach one subject (e.g., Chinese language). Still, there is one head teacher that is responsible for the daily affairs of the class (Chen, Rubin, & Li, 1997). As the responsibilities of these head teachers are most comparable to those of Dutch teachers, they were invited to participate in the present study.

During the school visits in March 2018, the first author applied a similar procedure for the data collection as in the Netherlands. The students' response rate was 99% and nonparticipation was mainly due to their absence. Instead of filling out a digital questionnaire, most of the Chinese teachers completed a paper questionnaire during the school visits. Two teachers, who were absent at the time of data collection, were emailed and completed a digital questionnaire within two weeks.

2.3. Measurements

2.3.1. Teachers' relationship perceptions

Teachers reported about the quality of their relationships with the eight selected students on the Dutch adapted version of the Student Teacher Relationship Scale (STRS, Pianta, 2001; Koomen, Verschueren, van Schooten, Jak, & Pianta, 2012). The STRS includes three dimensions to measure teachers' affective relationships with individual students: Closeness, Conflict, and Dependency. The Closeness subscale taps the degree of warmth, trust and open communication in the relationship, for instance *This child openly shares his/her feelings and experiences with me*. The Conflict subscale measures the degree of anger, negativity and unpredictability in the relationship, for example *This child and I always seems to be struggling with each other*. The Dependency subscale measures the extent to which the teacher experiences clingy and demanding behaviors from the child, for example *This child reacts strongly to separation from me*. Teachers indicated the degree to which each of the statements applied to their relationships with a particular child on a 5-point Likert scale, from 1 (*Definitely does not apply*) to 5 (*Definitely applies*).

In the present study, a short version of the Dutch adaptation of the STRS was used, which includes five items per dimension with the highest factor loadings reported in previous research (Koomen et al., 2012). For the Chinese sample, we selected items from the existing Chinese translation of the STRS that correspond to the same items as in the short Dutch version (Zhang, 2010). As one item

from the Closeness subscale and two items from the Dependency subscale were not available in the Chinese version, we translated them into Chinese with a back-translation procedure to examine the precision of wording.

The short version of the Dutch adapted STRS has showed satisfactory reliability in previous studies (Cronbach's alphas ranging from 0.85 to 0.86 for Closeness, 0.80 to 0.93 for Conflict, and being 0.91 for Dependency; Zee, de Jong, & Koomen, 2017; Zee & Koomen, 2017; Zee, Koomen, & van der Veen, 2013). The construct validity of the short version of the STRS has also been supported in previous research (Zee et al., 2013; Zee & Koomen, 2017). The complete Chinese version of the STRS showed satisfactory reliability too (Cronbach's alphas of 0.81 for Closeness, 0.83 for Conflict, and 0.67 for Dependency; Zhang, 2010), and the factorial validity and predictive validity was supported through confirmatory factor analysis and correlations with children's problem behaviors (Zhang, 2010). In the present study, the short STRS showed high reliability in the Dutch sample, with Cronbach's alphas of 0.86, 0.89, and 0.82 for Closeness, Conflict, and Dependency, respectively. For the Chinese sample, alphas were adequate to high, that is 0.83, 0.67, and 0.78 for Closeness, Conflict, and Dependency, respectively.

2.3.2. Students' relationship perceptions

Students' relationship perceptions were measured with two subscales from the Student Perception of Affective Relationship with Teacher Scale (SPARTS; Koomen & Jellesma, 2015), namely Closeness and Conflict. The Closeness subscale consists of eight items and taps students' positive feelings and reliance on teachers when they feel stressed, for example *I feel relaxed with my teacher*. The Conflict subscale includes ten items and describes students' perceptions of negative behaviors and distrustful feelings from teachers, for example *I easily have quarrels with my teacher*. Students rated the extent to which these statements applied to their relationships with the teacher on a 5-point Likert scale, ranging from 1 (*No, that is not true*) to 5 (*Yes, that is true*). As Dependency was not available as a subscale of the SPARTS (Koomen & Jellesma, 2015), we did not include this relationship dimension in students' relationship perceptions.

The Closeness (Cronbach's alpha ranging from 0.74 to 0.80) and Conflict subscale of the SPARTS (Cronbach's alpha ranging from 0.79 to 0.89) showed adequate reliability in previous studies (Jellesma et al., 2015; Koomen & Jellesma, 2015; Zee & de Bree, 2017; Zee & Koomen, 2017). The factorial, convergent, and concurrent validity of the SPARTS have also been supported in previous research (Koomen & Jellesma, 2015; Zee & de Bree, 2017). The SPARTS was originally developed in Dutch. To measure the perceptions of the Chinese students, the English version was translated into Chinese by the first author with a back translation procedure. In the present study, the Dutch version of the Closeness ($\alpha = 0.84$) and Conflict ($\alpha = 0.84$) subscales showed satisfactory reliability. In the Chinese sample, the Closeness ($\alpha = 0.83$) and Conflict ($\alpha = 0.72$) subscales also showed adequate to high reliability.

2.4. Analyses

Analyses were conducted with Mplus version 7 (Muthén & Muthén, 1998–2012). To answer the research questions, structural equation modeling (SEM) was used with three steps. First, we built multigroup models to test measurement invariance across countries for both teachers' and students' relationship perceptions. Second, we compared the latent means of teachers' and students' relationship perceptions between the two countries to investigate cultural differences in the relationship quality.² When comparing means between groups, a latent model is preferred above a comparison of observed variables in case of partial measurement invariance (Steinmetz, 2013). More specifically, using observed means may lead to either spurious differences (i.e., there are significant group differences in observed means although the groups have equal latent means) or compensation effect (i.e., despite significant differences in latent means, there is no significant difference in observed means; Steinmetz, 2013). As previous cross-cultural studies often found evidence for partial measurement invariance (Chen, 2008; and we did as well, see Results section), we decided to test group differences in latent means instead of observed means. Still, we did look into differences in observed means and provided results in Table 1. In line with Steinmetz (2013), we found evidence for both spurious differences and compensation effect, indicating that observed means may not be trustworthy in the case of partially strong invariance. Third, we estimated the correlations between teacher and student reports and examined the cultural differences in informant agreement. The details of each step are described in the following paragraphs.

2.4.1. Measurement invariance

Separate measurement models were built for teachers' and students' perceptions. According to Wolf, Harrington, Clark, and Miller (2013), at least 200 cases are needed for a two-factor model with missing values. The number of teachers in the Chinese sample (112 teacher reports in China) seems to be a bit small to allow fitting a three-factor model with 15 indicators.³ Therefore, we fitted three

² For all the SEM models, we also tested models including students' gender, grade, and ethnicity as covariates, since previous studies found these student characteristics to be associated with the quality of TSRs (see McGrath & van Bergen, 2015 for a review). However, mean differences in both students' and teachers' relationship perceptions revealed in the models including covariates were similar to the mean differences revealed in models without covariates (see Table 1 for mean differences revealed in models with and without covariates). Due to model parsimony considerations, we mainly reported results from the models without covariates.

³ For teachers' relationship perceptions, we also tested model fit of a three-factor model and a two-factor model with Closeness and Conflict. Both the three-factor model (RMSEA = 0.095, CFI = 0.852, SRMR = 0.099) and two-factor model (RMSEA = 0.108, CFI = 0.874, SRMR = 0.101), had a poor model fit. In addition, fitting separate models also increase statistical power to detect potential mean differences. We conducted a power analysis to calculate the power to detect a mean difference of 0.5 in a multigroup model with five observed variables and one common factor. We followed the steps in Muthén and Muthén (2002) and conducted a Monte Carlo simulation with 10,000 replications. The results revealed that given

Table 1

Observed mean scores and factor mean differences in teachers' and students' relationship perceptions.

	Observed mean scores				Factor mean differences from latent models					
	Mean(SD)		<i>t</i>	<i>p</i>	Model without covariates			Model with covariates		
	Dutch	Chinese			Diff (Std.)	<i>Z</i>	<i>p</i>	Diff (Std.)	<i>Z</i>	<i>p</i>
Teacher perceptions										
Closeness	3.96 (0.78)	3.64 (0.85)	3.38	< 0.001	−0.19 (−0.18)	−0.63	0.527	−0.17 (−0.20)	−0.62	0.535
Conflict	1.67 (0.91)	1.67 (0.61)	−0.04	0.970	−0.31 (−0.56)	−2.20	0.028	−0.29 (−0.53)	−1.98	0.048
Dependency	1.90 (0.84)	2.33 (0.79)	−4.58	< 0.001	0.39 (0.40)	1.34	0.180	0.40 (0.41)	1.34	0.180
Student perceptions										
Closeness	3.50 (0.86)	4.04 (0.84)	−11.57	< 0.001	0.72 (0.72)	6.95	< 0.001	0.67 (0.68)	6.53	< 0.001
Conflict	1.70 (0.67)	1.55 (0.52)	4.36	< 0.001	−0.31 (−0.40)	−2.68	0.007	−0.26 (0.33)	−3.44	0.001

Note. Observed means were calculated by computing the mean scores of the observed items measuring teachers' and students' relationship perceptions. With regards to latent models, students' gender, grade and ethnicity were included in the models with covariates. Diff refers to factor mean differences between Chinese and Dutch groups, taking Dutch group as the reference group. Std. refers to standardized coefficients of factor mean differences. Comparing group differences in observed mean scores to group differences in latent means, we found spurious differences for teacher-reported Closeness and Dependency, and a compensation effect for teacher-reported Conflict. These findings again addressed the importance of establishing adequate measurement invariance in cross-cultural studies.

separate models for the three relationship dimensions respectively, namely Closeness, Conflict and Dependency. For students' relationship perceptions, given the large students sample (789 in the Netherlands and 587 in China), the proposed measurement model consisted of two latent factors, Closeness and Conflict. We first build a baseline model (configural invariance) where all factor loadings and intercepts were freely estimated. To identify the model, all factor means were fixed to zero and all factor variances were fixed to one. To test weak invariance, we followed the steps suggested by Yoon and Millsap (2007). More specifically, we constrained the factor loadings to be equal across groups, fixed the variance(s) of the common factor(s) in the Dutch group to one and freely estimated the variance(s) of the common factor(s) in the Chinese group. To test strong invariance, similarly, we added equality constraints across groups to the latent item intercepts, fixed the means(s) of the common factors in the Dutch group to zero and freely estimated the mean(s) of the common factor(s) in the Chinese group.

2.4.2. Cultural differences in relationship quality

By means of the models with (partial) strong invariance, we could test the differences in latent means between the Dutch and Chinese samples. *Z* tests were used to examine whether mean differences between the Dutch and Chinese samples were different from zero. We standardized the latent means to ease interpretation of the magnitude of the mean differences. Standardized scores of 0.2, 0.5 and 0.8 were considered as cut-off values for small, medium, and large effects, respectively (Cohen, 1988).

2.4.3. Cultural differences in agreement

Based on the models with (partial) strong invariance, we extracted the factor scores and estimated the correlations between teachers' and students' perceptions in the Dutch and Chinese samples. Because Dependency was not included in the students' perceptions, only agreement in Closeness and Conflict was considered. To compare correlation coefficients, we employed the Fisher's *r* to *z* transformation (Fisher, 1915; Fisher, 1921), and used a two-tailed *Z*-test to examine cultural differences in teacher-student agreement.

2.4.4. Model estimation and model comparison

Full maximum likelihood estimation was used in all models to deal with missing values (9% of the students and 2% of the teachers missed information about at least one of the questions). To account for the nested structure of our data (i.e., students were nested within teachers), we used 'Type = Complex' option in Mplus, which provides adjusted standard errors and chi-square values taking into account non-independence of observations due to cluster sampling (Muthén & Muthén, 1998–2012).

Maximum likelihood with robust standard errors and chi-square (MLR) was applied to estimate the models. As chi-square tests are sensitive to sample size and may reject models with even trivial misfit (Chen, 2007), three additional fit indices, Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI) and Standardized Root Mean Square Residual (SRMR), were also included to indicate the overall goodness-of-fit. RMSEA values lower than 0.08 (Browne & Cudeck, 1992; Hu & Bentler, 1999), CFI values higher than 0.90 (Bentler, 1992), and SRMR values lower than 0.08 (Hu & Bentler, 1999) indicate satisfactory fit. To compare nested models, we used chi-square difference test and employed the Satorra–Bentler correction (Satorra & Bentler, 2010), since MLR was used as an estimation method. As chi-square difference tests suffer from the same drawback as the absolute chi-square tests

(footnote continued)

our teacher sample, the power to detect such an effect is 74.2%. Fitting a three-factor or a two-factor model would further harm the power. Based on the above information, we decided to fit three separate models for teachers' relationship perceptions.

Table 2

Model fit and model comparison statistics for multiple group models of teachers' and students' perceptions of TSRs.

	χ^2	df	p	RMSEA	90% CI	CFI	SRMR	$\Delta\chi^2$	df	p	Δ CFI	Δ RMSEA	Δ SRMR
Teachers' perceptions - closeness													
Configural invariance	33.32	10	< 0.001	0.111	[0.071, 0.153]	0.950	0.040						
Configural invariance.2	12.72	8	0.122	0.056	[0.000, 0.110]	0.990	0.022						
Weak invariance	15.62	12	0.209	0.040	[0.000, 0.089]	0.992	0.063	3.71	4	0.446	−0.002	−0.016	0.041
Strong invariance	75.25	16	< 0.001	0.139	[0.109, 0.172]	0.873	0.133	52.09	4	< 0.001	0.119	0.099	0.070
Partially strong invariance	16.68	13	0.214	0.039	[0.000, 0.086]	0.992	0.067	1.10	1	0.295	0.000	0.001	0.010
Teachers' perception - conflict													
Configural invariance	17.49	10	0.064	0.063	[0.000, 0.110]	0.979	0.029						
Weak invariance	25.14	14	0.033	0.065	[0.018, 0.105]	0.969	0.060	7.45	4	0.114	0.010	0.010	0.030
Strong invariance	49.28	18	< 0.001	0.096	[0.064, 0.128]	0.914	0.111	19.14	4	0.001	0.055	0.030	0.020
Partially strong invariance	23.88	16	0.092	0.051	[0.000, 0.091]	0.978	0.060	0.45	2	0.800	−0.009	0.019	0.002
Teachers' perception - dependency													
Configural invariance	45.84	10	< 0.001	0.137	[0.098, 0.179]	0.901	0.045						
Configural invariance.2	18.70	8	0.017	0.084	[0.034, 0.134]	0.970	0.035						
Weak invariance	23.13	12	0.028	0.070	[0.023, 0.112]	0.969	0.060	5.43	4	0.246	0.001	−0.014	0.025
Strong invariance	43.26	16	< 0.001	0.095	[0.061, 0.129]	0.925	0.114	16.73	4	0.002	0.044	0.025	0.054
Partially Strong invariance	26.43	14	0.023	0.068	[0.025, 0.108]	0.966	0.070	3.32	2	0.190	0.003	−0.002	0.010
Students' perceptions													
Configural invariance	566.66	268	< 0.001	0.040	[0.036, 0.045]	0.925	0.045						
Weak invariance	612.34	284	< 0.001	0.041	[0.037, 0.046]	0.918	0.056	41.56	16	< 0.001	0.007	0.001	0.011
Strong invariance	132.25	300	< 0.001	0.071	[0.067, 0.074]	0.744	0.088	4007.23	16	< 0.001	0.174	0.028	0.029
Partially strong invariance	629.87	292	< 0.001	0.041	[0.037, 0.046]	0.915	0.056	17.61	8	0.024	0.003	0.000	0.000

Note. All the models with (partially) strong invariance were compared to the model with weak invariance.

(Chen, 2007), CFI changes (Δ CFI), RMSEA changes (Δ RMSEA) and SRMR changes (Δ SRMR) were included as alternative indicators. Model equivalence was indicated by nonsignificant Chi-square differences, Δ CFI values ≤ 0.010 (Cheung & Rensvold, 1999; French & Finch, 2006; Rutkowski & Svetina, 2014), Δ RMSEA values ≤ 0.010 (Chen, 2007; Rutkowski & Svetina, 2014), and Δ SRMR values ≤ 0.030 for weak invariance and ≤ 0.010 for strong invariance (Chen, 2007). When either chi-square (difference) tests or the three fit indices indicated satisfactory fit or model equivalence, we proceeded to the next step (Little, 2013).

In case measurement invariance did not hold, we checked standardized mean residuals in both groups and modification indices (MI) to locate the items showing non-invariance (Bentler, 1992; Bentler & Chou, 1993), and thereafter released the equality constraints on these items. This resulted in models with partial measurement invariance, which still allow valid comparisons between latent means, as long as at least two factor loadings and intercepts can be constrained to be equal across groups (Byrne, Shavelson, & Muthén, 1989).

3. Results

3.1. Measurement invariance in teachers' perceptions

3.1.1. The Closeness model

The model fit and model comparison statistics of multigroup models for teachers' perceptions are provided in Table 2. The fit of the measurement model for Closeness was not optimal, $\chi^2(10) = 33.32, p < .001$, RMSEA = 0.111, CFI = 0.950, SRMR = 0.040. To improve the model, we checked the correlation residuals and MI, and decided to add residual correlations between item 1 (*This child openly shares his/her feelings and experiences with me*) and item 3 (*If upset, this child will seek comfort from me*) because these two items both tapped a child's specific behavior in seeking open communications with the teacher. The adjusted model fitted the data satisfactorily, $\chi^2(8) = 12.72, p = .122$, RMSEA = 0.056, CFI = 0.990, SRMR = 0.022. The weak invariance model fitted the data equally well as the adjusted measurement model, $\Delta\chi^2(4) = 3.71, p = .446$, Δ CFI = −0.002, Δ RMSEA = −0.016, Δ SRMR = 0.041. The strong invariance model, however, fitted significantly worse than the model with weak invariance, $\Delta\chi^2(4) = 52.09, p < .001$, Δ CFI = 0.119, Δ RMSEA = 0.099, Δ SRMR = 0.070. Checking the mean residuals and MI, we released the equality constraints on the intercepts of item 9, 12 and 15 (see Appendix A). Item 9 and 12 had a higher intercept in the Dutch group, indicating that given the same level of Closeness, Dutch teachers scored higher on these items. On the contrary, item 15 had a higher intercept in the Chinese group, indicating that given the same level of Closeness, Chinese teachers scored higher on this item. The model with partially strong invariance reached good fit, $\chi^2(13) = 16.68, p = .214$, RMSEA = 0.039, CFI = 0.992, SRMR = 0.067, and fitted the data equally well as the model with weak invariance, $\Delta\chi^2(1) = 1.10, p = .295$, Δ CFI = 0.000, Δ RMSEA = 0.001, Δ SRMR = 0.010. Therefore, the model with partially strong invariance was used as the final model for further analyses (Byrne et al., 1989; see Fig. 1 for the coefficients).

3.1.2. The conflict model

The measurement model for Conflict had a satisfactory fit, $\chi^2(10) = 17.49, p = .064$, RMSEA = 0.063, CFI = 0.979,

Fig. 1. The estimated coefficients of partially strong invariance models for teachers' perceptions.
Note. *: $p < .05$. The coefficients between brackets are standardized coefficient.

SRMR = 0.029. The model with weak invariance also fitted the data equally well, $\Delta\chi^2(4) = 7.45$, $p = .114$, $\Delta CFI = 0.010$, $\Delta RMSEA = 0.010$, $\Delta SRMR = 0.030$. However, the model with strong invariance fitted the data significantly worse than the model with weak invariance, $\Delta\chi^2(4) = 19.14$, $p < .001$, $\Delta CFI = 0.055$, $\Delta RMSEA = 0.030$, $\Delta SRMR = 0.020$. Checking the mean residuals and MI, we released the constraints on the intercepts of item 7 and 11 (see Appendix A). The model with partially strong invariance had satisfactory fit, $\chi^2(16) = 23.88$, $p = .092$, $RMSEA = 0.051$, $CFI = 0.978$, $SRMR = 0.060$, and fitted the data equally well as the model with weak invariance, $\Delta\chi^2(2) = 0.45$, $p = .800$, $\Delta CFI = -0.009$, $\Delta RMSEA = 0.019$, $\Delta SRMR = 0.002$. The model with partially strong invariance was taken as the final model for further analyses (see Fig. 1 for the coefficients).

3.1.3. The dependency model

The fit of the measurement model for Dependency was not optimal, $\chi^2(10) = 45.84$, $p < .001$, $RMSEA = 0.137$, $CFI = 0.901$, $SRMR = 0.045$. Checking the correlation residuals and MI, we decided to add a residual correlation between item 8 (*This child reacts strongly to separation from me*) and 10 (*This child is overly dependent on me*), as these items both describe a child's response to an increased personal space with the teacher. Adding the residual correlation led to an improvement of model fit, $\chi^2(8) = 18.70$, $p = .017$, $RMSEA = 0.084$, $CFI = 0.970$, $SRMR = 0.035$. The model with weak invariance fitted the data equally well as the adjusted measurement model, $\Delta\chi^2(4) = 5.43$, $p = .246$, $\Delta CFI = 0.001$, $\Delta RMSEA = -0.014$, $\Delta SRMR = 0.025$. However, the model with strong invariance fitted the data significantly worse than the model with weak invariance, $\Delta\chi^2(4) = 16.73$, $p = .002$, $\Delta CFI = 0.044$, $\Delta RMSEA = 0.025$, $\Delta SRMR = 0.054$. Checking the mean residuals and MI, we decided to release equality constraints on the intercepts of item 4 and 6 (see Appendix A). The model with partially strong invariance reached satisfactory fit, $\chi^2(14) = 26.43$, $p = .023$, $RMSEA = 0.068$, $CFI = 0.966$, $SRMR = 0.070$, and the model fit did not deteriorate significantly from the model with weak invariance, $\Delta\chi^2(2) = 3.32$, $p = .190$, $\Delta CFI = 0.003$, $\Delta RMSEA = -0.002$, $\Delta SRMR = 0.010$, so it was taken as the final model for further analyses (see Fig. 1 for coefficients).

3.2. Measurement invariance in students' perceptions

The model fit and model comparison statistics of multigroup models for students' perceptions are also displayed in Table 2. The measurement model for students' perceptions had a satisfactory fit, $\chi^2(268) = 566.66$, $p < .001$, $RMSEA = 0.040$, $CFI = 0.925$, $SRMR = 0.045$. The model with weak invariance did not fit significantly worse than the measurement model, $\Delta\chi^2(16) = 41.56$, $p < .001$, $\Delta CFI = 0.007$, $\Delta RMSEA = 0.001$, $\Delta SRMR = 0.011$. However, the model with strong invariance fitted the data significantly worse than the model with weak invariance, $\Delta\chi^2(16) = 4007.23$, $p < .001$, $\Delta CFI = 0.174$, $\Delta RMSEA = 0.028$, $\Delta SRMR = 0.029$. Checking the standardized mean residuals and MI, we decided to free the equality constraints on the intercepts of Closeness items 1, 3, 5, 10, 15, 20, and Conflict items 8 and 23 (see Appendix A). Freeing these parameters resulted in a model with partially strong invariance, and this model fitted the data equally well as the model with weak invariance, $\Delta\chi^2(8) = 17.61$, $p = .024$, $\Delta CFI = 0.003$, $\Delta RMSEA = 0.000$, $\Delta SRMR = 0.000$. Therefore, partially strong invariance did hold for students' relationship perceptions and this model was taken as the final model (see Fig. 2 for the estimated coefficients).

To summarize, all three subscales of the STRS (Closeness, Conflict, and Dependency) and the two subscales of the SPARTS (Closeness and Conflict) reached partially strong measurement invariance across countries, with some items showing non-invariance in intercepts. These results indicated that most of the items measured the same constructs across two countries, whereas some items may measure different constructs or have different connotations in the Netherlands and China. According to Byrne et al. (1989), as well as Steenkamp and Baumgartner (1998), meaningful comparison between groups can still be made under the condition of partially strong invariance. Therefore, we proceed to the next step.

3.3. Cross-cultural differences in relationship quality

3.3.1. Teachers' perceptions

Based on the final models for teachers' relationship perceptions, we tested the mean differences of teacher-reported Closeness, Conflict, and Dependency across the two countries. For Closeness, the mean difference was -0.19 , which was not significantly different from zero, $Z = -0.63$, $p = .527$, 95% CI = $[-0.81, 0.42]$. In terms of Conflict, the mean difference was -0.31 , which was significantly different from zero, $Z = -2.20$, $p = .012$, 95% CI = $[-0.58, -0.03]$, with a medium effect size (Cohen's $d = -0.56$). With regard to Dependency, the mean difference was 0.39 , which did not differ significantly from zero, $Z = 1.34$, $p = .180$, 95% CI = $[-0.18, 0.96]$. These results indicated that compared to Dutch teachers, Chinese teachers perceived lower levels of Conflict and comparable levels of Closeness and Dependency in the relationships with their students.

3.3.2. Students' perceptions

Based on the final model for students' relationship perceptions, we tested the mean differences of student-reported Closeness and Conflict across the Netherlands and China. For Closeness, the mean difference between the Chinese and Dutch students was 0.71 , which was significantly different from zero, $Z = 6.95$, $p < .001$, 95% CI = $[0.51, 0.91]$, with a medium effect size (Cohen's $d = 0.72$). In terms of Conflict, the mean difference was -0.31 , which was significantly different from zero, $Z = -2.68$, $p = .007$, 95% CI = $[-0.46, -0.16]$, with a small effect size (Cohen's $d = -0.40$). These results implied that Chinese students experienced higher levels of Closeness and lower levels of Conflict with teachers compared to Dutch students.

Fig. 2. The estimated coefficients of the partially strong invariance model for students' perceptions. Note. ***: $p < .001$. The coefficients between brackets were standardized coefficients.

Table 3

The correlations between the students' and teachers' perceptions towards closeness and conflict in the Netherlands and China.

	Closeness-T	Conflict-T	Closeness-S	Conflict-S
Closeness-T	–	0.06	0.36***	–0.24*
Conflict-T	–0.40***	–	–0.11	0.17
Closeness-S	0.15*	–0.32***	–	–0.73***
Conflict-S	–0.17**	0.42***	–0.83***	–

Note. Closeness-T and Conflict-T refer to the teachers' perceptions while Closeness-S and Conflict-S refer to the students' perceptions. The correlations in the Dutch sample are presented in the lower-left corner and the correlations in the Chinese sample are presented in the upper-right corner.

*** $p < .001$.

** $p < .01$.

* $p < .05$.

3.4. Agreement in student and teacher perceptions across cultures

The estimated Pearson correlations between teachers' and students' perceptions of Closeness and Conflict are provided in Table 3. As can be seen in this table, the correlations for agreement between Dutch teacher-student dyads were significant and positive for both Closeness ($r = 0.15$, $p = .011$) and Conflict ($r = 0.42$, $p < .001$). With regard to the agreement in Chinese teacher-student dyads, the correlation for Closeness ($r = 0.36$, $p < .001$) was also significant and positive. Surprisingly, however, the correlation between students' and teachers' perceptions of Conflict was not significant ($r = 0.17$, $p = .075$). The correlation between teacher-reported and student-reported Closeness was significantly higher in the Chinese group ($r = 0.36$) than the correlation in the Dutch group ($r = 0.15$), $Z = 1.98$, $p = .044$. On the contrary, the correlation between teacher-reported and student-reported Conflict was significantly higher in the Dutch group ($r = 0.42$) compared to the Chinese group ($r = 0.17$), $Z = -2.43$, $p = .015$. These results indicated that, compared to Dutch teacher-student dyads, Chinese teacher-student dyads had higher agreement on the level of closeness but lower agreement on the level of conflict in their mutual relationships.

4. Discussion

In this study, we investigated the affective quality of teacher-student relationships in the Netherlands, a Western, individualistic country, and in China, an Eastern, collectivistic country. Three main findings emerged from the present study: First, we found evidence for partially strong measurement invariance for both teachers' and students' relationship perceptions. Second, both Chinese teachers and students reported lower levels of conflict than Dutch teachers and students. Chinese students also experienced more closeness in their relationships with teachers than Dutch students, whereas this difference was not found in teachers' perceptions. Third, Chinese teacher-student dyads had higher agreement in reporting closeness, whereas Dutch teacher-student dyads agreed more in their perceptions of conflict.

4.1. Partially strong invariance in relationship perceptions

We found partially strong invariance for the SPARTS and STRS, indicating that some items may have different connotations in the Netherlands and China. Similarly, Jia et al. (2009) and Yang et al. (2013) also found only partially strong invariance with the instruments they employed (i.e., a revised school climate measure and the Modified-Delaware School Climate Survey-Student) in their comparisons between American and Chinese students. These findings are in line with the literature review by Chen (2008), that it is common for cross-cultural studies to find only partial measurement equivalence. Given partial measurement invariance, we still proceed to test mean differences because of three reasons. First, according to Byrne et al. (1989), as well as Steenkamp and Baumgartner (1998), meaningful comparison between groups can still be made, as long as at least two items are invariant. Second, statistical bias in group comparison is dramatically reduced to an appreciable level in case the non-invariant items are balanced across groups (i.e., some of the items are higher in one group whereas others are higher in another group; Chen, 2008). This applies to the distribution of invariant items we found in the SPARTS and STRS (see Appendix A for details), indicating that our findings may not be much biased by the non-invariant items. Third, as suggested by both Chen (2008) and Putnick and Bornstein (2016), we compared the mean differences (research question 2) revealed in the models with and without imposing equality constraints on non-invariant items. As the results turned out to be similar, statistically it may be justifiable for us to make group comparisons (Chen, 2008).

Nevertheless, conceptually some items still convey different meanings in different countries (Chen, 2008). Therefore it is important to look into the item contents and explore the possible reasons for these different meanings. With regard to teachers' perceptions, for example, given the same level of conflict with students, Chinese teachers had a higher score than Dutch teachers on the item *When this child is in a bad mood, I know we are in for a long and difficult day*. This might be due to the emphasis on children's respect and obedience to teachers in collectivistic cultures (Bear et al., 2014; Triandis, 2018; Yang et al., 2013). Chinese teachers might therefore be more critical towards children's deviant behaviors in teacher-student interactions (Yang et al., 2013). With regard to students' perceptions, for instance, given the same level of closeness and conflict, Dutch students had a higher score than Chinese students on the items *If I have a problem, I can share it with my teacher* and *I can be very angry with my teacher*. A plausible explanation is

the different expectations for interpersonal interactions that individuals hold in collectivistic and individualistic cultures. In collectivistic cultures, individuals are expected to suppress feelings and emotions in interpersonal relationships (Matsumoto et al., 2008; Rubin et al., 2006; Wu et al., 2002) whereas in individualistic cultures, individuals are encouraged to express their feelings and emotions openly and freely (Matsumoto et al., 2008). Therefore, Chinese students may be less inclined to share their problems and anger with teachers. As an implication, researchers and practitioners ought to be more sensitive and cautious when applying the instruments developed in Western contexts to measure the quality of TSRs in Eastern contexts. Future research is also needed to develop culturally sensitive instruments that take into account the different understandings of relationship quality in different cultures.

In addition, for both teachers' and students' perceptions, the closeness dimension contained most items varying across the Netherlands and China, indicating that the meaning of closeness might be more likely to differ across collectivistic and individualistic cultures, whereas the connotations of conflict and dependency are relatively similar. Future studies may as well further explore plausible explanations about why the closeness dimension seems to differ most across cultures, probably by using in-depth interviews with teachers and students from both cultures.

4.2. Cross-cultural differences in relationship quality

In line with our hypothesis, we found that Chinese students experienced higher levels of closeness with their teachers than their Dutch counterparts. This finding is consistent with findings from Jia et al. (2009) and Yang et al. (2013), who found that students in China perceived more teacher support/closeness compared to American students. Together, these findings may imply that students in collectivistic cultures tend to experience more closeness in their relationship with teachers than students in individualistic cultures, which might be due to the focus on interpersonal connectedness in collectivistic cultures (Markus & Kitayama, 1991).

In contrast with our expectations (Beyazkurk & Kesner, 2005; Jia et al., 2009; Yang et al., 2013), however, Chinese and Dutch teachers appeared to hold equally favorable perceptions of the degree of closeness in their relationships with students. This finding is interesting considering the group difference we found in students' perception of closeness in TSRs. However, mixed results were found in two previous studies including teachers' relationship perceptions. That is, Turkish kindergarten teachers reported more closeness than their American counterparts (Beyazkurk & Kesner, 2005), whereas American elementary school teachers reported higher teacher support than their Chinese counterparts (Bear et al., 2014). Given these findings, it remains unclear whether there are, in general, cross-cultural differences in teachers' perception of closeness. Therefore, future studies should further look into it.

As expected (Dennis et al., 2002), both Chinese students and teachers perceived less conflict in their mutual relationships compared to their Dutch counterparts. Found in both teachers' and student's relationship perceptions, the cultural differences in conflict are more salient than the differences in closeness. A plausible reason for these findings might be the difference in the power distance between teachers and students in the Netherlands and China. More specifically, the traditional Chinese culture puts an emphasis on children's compliance and obedience to parents and teachers, potentially resulting in a large power distance between students and teachers (Bear et al., 2014; Yang et al., 2013). As a result, students are expected to respect their teachers and avoid deviant behaviors that might raise conflict (Bear et al., 2014; Yang et al., 2013). On the contrary, Dutch culture often emphasizes an egalitarian point of view with a lower power distance between teachers and students (den Brok & Levy, 2006; Hofstede et al., 2010). As a result, teachers grant students a considerable freedom of action and students are more critical of the curriculum and teaching methods (Boekaerts, 2003). Therefore, compared to Dutch students, Chinese students may display more socially desirable behaviors (Yang et al., 2013), and be less critical about their teachers (Jia et al., 2009), which might in turn elicit less observable conflict with teachers.

Different from our expectations (French et al., 2001), there was no significant cultural difference in teachers' perceptions of the degree of dependency in their relationships with students. A plausible reason might be that cultural values affect both students' dependent behaviors and teachers' understanding about dependency. For example, Chinese students may actually display more dependent behaviors than Dutch students. Yet, the emphasis of interdependence in collectivistic cultures (i.e. demanding more dependency from students; Markus & Kitayama, 1991; Triandis, 2018) may also increase Chinese teachers' tolerance for these behaviors and, hence, Chinese teachers will not perceive students as being more dependent. Another plausible explanation is that the larger class size in China (around 40 students per class versus around 20 students per class in the Netherlands) makes it less likely for students to be overly dependent on teachers and, therefore, no cultural difference in dependency was found.

Altogether, our findings seem to suggest that there may be cultural differences in teachers' and especially in students' relationship perceptions. Therefore, findings drawn from Western samples may not be simply generalizable to Eastern samples. As an implication, more studies including Eastern countries are needed to investigate how research findings from Western studies (e.g., associations between TSRs and students' social emotional and academic development; Lei et al., 2016; Roorda et al., 2017) may transfer to Eastern samples. Furthermore, more research is needed to find out whether relationship-focused interventions developed in Western countries (e.g., Driscoll & Pianta, 2010; Spilt, Koomen, Thijs, & van der Leij, 2012) would also be effective for Eastern teachers and students.

4.3. Cross-cultural differences in informant agreement

Due to a lack of empirical research, we did not have a hypothesis about cultural differences in informant agreement. Our results revealed that Chinese teacher-student dyads displayed higher agreement in perceptions with regard to closeness and lower agreement with respect to conflict, compared to Dutch dyads. A plausible reason for the higher agreement on closeness in Chinese teacher-

student dyads might be the emphasis on social harmony in collectivistic cultures (Chen & French, 2008; Yang et al., 2013). Maintaining harmonious social interactions requires a full understanding of how others are feeling and thinking in the relationship (Markus & Kitayama, 1991). Hence, teacher-student dyads in collectivistic cultures might be more focused and more sensitive about each other's perceptions, leading to more consensus about the degree of closeness in their mutual relationship.

For the cultural difference in informant agreement about the degree of conflict in the relationship, a plausible reason might be the preferred conflict-solving strategies in both cultures. Individuals in collectivistic cultures tend to avoid bringing up possible conflicts in the relationships (Triandis et al., 1988), which in turn may make it more difficult for Chinese teacher-student dyads to learn about each other's dissatisfaction and struggles in the relationship. For example, a Chinese teacher may never know if students felt unfairly treated as they would not say it out loud. In line with this explanation, we found that the correlation between teacher-reported and student-reported conflict was not significant in the Chinese sample. In contrast, teacher-student dyads from Western countries like the Netherlands may be more inclined to express their dissatisfaction with the relationship and to talk about conflicts freely (den Brok & Levy, 2006; Triandis, 2018). This is confirmed by the significant correlation between teacher-reported and child-reported conflict we found in the Dutch sample. These findings suggest that cultural values might cast influences on the way teachers and student interact with each other and thereafter the agreement in teachers' and students' relationship perceptions. However, as the present study is, as far as we know, the first to examine cultural differences in agreement between teachers' and students' relationship perceptions, more research is needed before more stringent conclusions can be drawn. Future studies may, for example, include in-depth interviews with teachers and students, as well as observations of real-time teacher-student interactions, to get more insight about why teachers and students have different relationship perceptions.

4.4. Practical implications

As findings of the present study are exploratory and preliminary, caution is warranted when generalizing our findings to other populations, for example, teachers and students from other countries. However, some practical implications can still be made. First, teachers and school psychologists may get inspiration to improve the quality of TSRs by looking into how teachers and students interact with each other in another cultural context. For example, given the lower closeness and higher conflict found between Dutch teachers and students, it may be beneficial for Dutch teachers to learn from their Chinese colleagues and attach more importance to connectedness and sensitivity in relationships. In addition, given the low teacher-student agreement on conflict for Chinese teacher-student dyads, Chinese teachers may benefit from their Dutch counterparts by being more open about their feelings and emotions in the relationship. These practices may provide students in both countries with a better context to comfortably explore the learning environment.

Second, due to the presence of potential cultural differences, it is essential for researchers, school psychologists, and teachers to be aware that interventions developed in one context may not be simply applied to another cultural context. As a result, it may be important to explore, for example, how relationship-focused interventions developed in Western countries (e.g., Driscoll & Pianta, 2010; Spilt et al., 2012) may be effectively adjusted in Eastern countries. For instance, the relationship-focused reflection intervention (Spilt et al., 2012) aims to help teachers reflecting on their strengths and difficulties in several dimensions such as sensitivity of discipline. However, as sensitivity might have different connotations in Western and Eastern cultures (Rothbaum et al., 2000), statements mentioned in this intervention (e.g., talk with the students openly about their misbehaviors) may not completely apply to Eastern countries like China. Therefore, it seems that adjustment is needed to make the interventions developed in Western countries more adaptive to Eastern countries.

Third, given that schools nowadays are becoming more and more multicultural, it may be important for teachers and school psychologists to be more culturally sensitive and increase their awareness about the influence of cultural values on the quality of TSRs. As the quality of TSRs is often found to associate with students' social emotional and academic development (Lei et al., 2016; Roorda et al., 2017), researchers and practitioners may also look into how cultural values may cast influence on the association between the relationship quality and students' learning and well-being. For instance, teachers may consciously reflect on how their common beliefs (e.g., there should/should not be a large distance between teachers and students) affect their interactions with students, and thereafter students' learning and social emotional behaviors.

4.5. Limitations and suggestions for future research

Some limitations need to be taken into account when interpreting the results of the present study. First, given the complex nature of cross-cultural designs, it is difficult to isolate 'genuine cultural differences' (i.e., group differences produced by cultural values) from differences produced by non-cultural factors (e.g., group difference produced by economic factors). Although we tested measurement invariance and took into account covariates such as students' gender, there may still be other factors confounding the group differences in the quality of TSRs, for example, students' socioeconomic status. Without further research, we cannot draw more definitive conclusions about whether there are true cultural differences in the affective quality of TSRs between collectivistic and individualistic cultures. Future studies are therefore encouraged to include more heterogeneous samples (e.g., teachers and students from other countries), as well as various research designs (e.g., priming) to identify the source of the group differences.

Second, although both teachers' and students' relationship perceptions were included in the present study, we did not focus on the third element of teacher-student relationships, namely, the real-time, observable teacher-student interactions (Pianta et al., 2003). Investigating real-time teacher-student interactions may further help addressing the mechanism of how cultural values cast an influence on the development of TSRs, based on which culturally designed interventions can be developed. Therefore, future research is

needed in which these interactions are included. Such studies may deepen our understanding of cultural difference in TSRs.

Third, given the smaller teacher sample size, three separate models had to be fitted for teachers' relationship perceptions. It would be interesting to fit three relationship dimensions in a single measurement model, where the covariance between relationship dimensions is also estimated. Future studies are therefore suggested to include a larger teacher sample to further look into the measurement invariance of the STRS in cross-cultural comparisons. In addition, the relatively small teacher sample may limit the generalizability of our findings to a larger teacher population (e.g., teachers with more diverse ethnicity backgrounds). Therefore, future research is also encouraged to include a larger and more heterogeneous teacher sample to investigate whether our findings are generalizable to other teachers.

5. Conclusion

When looking into measurement invariance in relationship perceptions of upper elementary school students and their teachers, both students' perceptions of closeness and conflict and teachers' perceptions of closeness, conflict, and dependency reached partially strong invariance across the Netherlands and China. Therefore, we conclude that, in general, our findings indicate that the SPARTS and STRS can be used in cross-cultural settings (cf. Jia et al., 2009; Yang et al., 2013). Still, caution is needed with regard to the items with non-invariant intercepts across cultures.

Furthermore, Chinese students in upper primary schools appeared to experience more closeness and less conflict in their relationships with teachers than their Dutch counterparts. Chinese teachers also experienced less conflict but not more closeness and dependency in their relationships with students as compared to Dutch teachers. Given the potential presence of cultural differences in the quality of TSRs, caution should be warranted when generalizing research findings (e.g., association between TSRs quality and students' school functioning), or applying relationship-focused interventions (e.g., Driscoll & Pianta, 2010; Spilt et al., 2012) developed in Western contexts to other contexts.

Finally, Chinese teacher-student dyads agreed more about the degree of closeness in their mutual relationships but agreed less about the degree of conflict than Dutch teacher-child dyads. Research including observations of real-time teacher–student interactions and in-depth interviews may help further explaining these findings. In sum, our study seems to confirm the theoretical assumption (Bowlby, 1980; Pianta et al., 2003) that differences in cultural values may transfer to different interpretations of teacher-student relationships. Therefore, findings drawn from one cultural context may not be simply generalized to another. However, future research is needed before more stringent conclusion can be drawn.

Acknowledgements

This work was supported by the CSC scholarship offered by the China Scholarship Council in collaboration with the University of Amsterdam.

Appendix A

Table A1
Components and measurement non-invariance of the teachers' perceptions of relationship quality.

Item	Content	Non-invariance
Closeness		
STRS1	This child openly shares his/her feelings and experiences with me.	/
STRS3	If upset, this child will seek comfort from me.	/
STRS9	I share an affectionate, warm relationship with this child	Dutch Sample
STRS12	This child seems to feel secure with me	Dutch Sample
STRS15	My interactions with this child make me feel effective and confident	Chinese Sample
Conflict		
STRS2	This child and I always seem to be struggling with each other.	/
STRS5	This child feels that I treat him/her unfairly.	/
STRS7	When this child is in a bad mood, I know we are in for a long and difficult day	Chinese Sample
STRS11	This child's feelings towards me can be unpredictable or can change suddenly	Chinese Sample
STRS13	Dealing with this child drains my energy.	/
Dependency		
STRS4	This child needs to be continually confirmed by me.	Chinese Sample
STRS6	This child fixes his/her attention on me the whole day long.	Dutch Sample
STRS8	This child reacts strongly to separation from me.	/
STRS10	This child is overly dependent on me.	/
STRS14	This child asks for my help when he/she really does not need help.	/

Note. The sample referred in the non-invariance column is the sample with higher intercept in the corresponding item. / stands for measurement invariance.

Table A2

Components and measurement non-invariance of the students' perceptions of relationship quality.

Item	Content	Non-invariance
Closeness		
SPARTS1	I feel at ease with my teacher	Dutch Sample
SPARTS3	When I don't feel well, my teacher notices and asks me about it	Chinese Sample
SPARTS5	When I feel uncomfortable, I go to my teacher for help and comfort	Chinese Sample
SPARTS10	I tell my teacher things that are important to me	Dutch Sample
SPARTS11	My teacher understands me	/
SPARTS15	I think I have a good relationship with my teacher	Dutch Sample
SPARTS20	If I have a problem, I can share it with my teacher	Dutch Sample
SPARTS25	My teacher usually knows how I feel	/
Conflict		
SPARTS4	Other children are less punished	/
SPARTS7	I easily have quarrels with my teacher	/
SPARTS8	My teacher particularly tells me what I do wrong and not what I do right	Chinese Sample
SPARTS12	My teacher treats me unfairly.	/
SPARTS14	I guess my teacher thinks I whine a lot.	/
SPARTS16	My teacher thinks I do things sneaky.	/
SPARTS19	I guess my teacher gets tired of me in class.	/
SPARTS21	I feel that my teacher does not trust me.	/
SPARTS23	I can be very angry with my teacher	Dutch Sample
SPARTS24	If the teacher says something is not allowed, I often do it anyway.	/

Note. The sample referred in the non-invariance column is the sample with higher intercept in the corresponding item. / stands for measurement invariance.

References

- Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment*. Hillsdale, NJ: Erlbaum.
- Aldridge, J. M., & Fraser, B. J. (2000). A cross-cultural study of classroom learning environments in Australia and Taiwan. *Learning Environments Research*, 3, 101–134. <https://doi.org/10.1023/A:1026599727439>.
- Bear, G. G., Yang, C., Glutting, J., Huang, X., He, X., Zhang, W., & Chen, D. (2014). Understanding teacher-student relationships, student-student relationships, and conduct problems in China and the United States. *International Journal of School and Educational Psychology*, 2, 247–260. <https://doi.org/10.1080/21683603.2014.883342>.
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the bulletin. *Psychological Bulletin*, 112, 400–404.
- Bentler, P. M., & Chou, C. P. (1993). Some new covariance structure model improvement statistics. *Sage Focus Editions*, 154, 235–255.
- Beyazkurk, D., & Kesner, J. E. (2005). Teacher-child relationship in Turkish and United States school: A cross-cultural study. *International Education Journal*, 6, 574–554. Retrieved from <http://iej.cjb.net>.
- Boekaerts, M. (2003). Adolescence in Dutch culture: A self-regulation perspective. In F. Pajares, & T. Urdan (Eds.). *Adolescence and education: Vol. 3. International perspectives on adolescence* (pp. 101–124). Greenwich, CT: Information Age.
- Bowlby, J. (1980). *Attachment and loss: Vol.3. Loss: Sadness and depression*. NewYork: BasicBooks.
- den Brok, P., & Levy, J. (2006). Teacher-student relationships in multicultural classes: Reviewing the past, preparing the future. *International Journal of Educational Research*, 43, 72–88. <https://doi.org/10.1016/j.ijer.2006.03.007>.
- Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21, 230–258. <https://doi.org/10.1177/0049124192021002005>.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456–466. <https://doi.org/10.1037/0033-2909.105.3.456>.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14, 464–504. <https://doi.org/10.1080/10705510701301834>.
- Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, 95, 1005–1018. <https://doi.org/10.1037/a0013193>.
- Chen, X., & French, D. C. (2008). Children's social competence in cultural context. *Annual Review of Psychology*, 59, 591–616. <https://doi.org/10.1146/annurev.psych.59.103006.093606>.
- Chen, X., Rubin, K. H., & Li, D. (1997). Relation between academic achievement and social adjustment: Evidence from Chinese children. *Developmental Psychology*, 33, 518–525.
- Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management*, 25, 1–27. <https://doi.org/10.1177/014920639902500101>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Dennis, T. A., Cole, P. M., Zahn-Waxler, C., & Mizuta, I. (2002). Self in context: Autonomy and relatedness in Japanese and U.S. mother-preschooler dyads. *Child Development*, 73, 1803–1817. <https://doi.org/10.1111/1467-8624.00507>.
- Driscoll, K. C., & Pianta, R. C. (2010). Banking time in head start: Early efficacy of an intervention designed to promote supportive teacher-child relationships. *Early Education and Development*, 21, 38–64. <https://doi.org/10.1080/10409280802657449>.
- Dwairy, M., & Achoui, M. (2010). Adolescents-family connectedness: A first cross-cultural research on parenting and psychological adjustment of children. *Journal of Child and Family Studies*, 19, 8–15. <https://doi.org/10.1007/s10826-009-9335-1>.
- Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10, 507–521. <https://doi.org/10.2307/2331838>.
- Fisher, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron*, 1, 1–32.
- French, B. F., & Finch, W. H. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 13, 378–402. https://doi.org/10.1207/s15328007sem1303_3.
- French, D. C., Rianasari, M., Pidada, S., Nelwan, P., & Buhrmester, D. (2001). Social support of Indonesian and U. S. children and adolescents by family members and friends. *Merrill-Palmer Quarterly*, 47, 377–394. <https://doi.org/10.1353/mpq.2001.0015>.

- Gregoriadis, A., & Grammatikopoulos, V. (2014). Teacher-child relationship quality in early childhood education: The importance of relationship patterns. *Early Child Development and Care*, 184, 386–402. <https://doi.org/10.1080/03004430.2013.790383>.
- Gregoriadis, A., & Tsigilis, N. (2008). Applicability of the student-teacher relationship scale (STRS) in the Greek educational setting. *Journal of Psychoeducational Assessment*, 26, 108–120. <https://doi.org/10.1177/0734282907306894>.
- Hofstede, G. (1980). Culture and organizations. *International Studies of Management & Organization*, 10, 15–41. <https://doi.org/10.1080/00208825.1980.11656300>.
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind* (Rev. 3rd ed.). New York: McGraw-Hill.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. <https://doi.org/10.1080/10705519909540118>.
- Hughes, J. N. (2011). Longitudinal effects of teacher and student perceptions of teacher-student relationship qualities on academic adjustment. *The Elementary School Journal*, 112, 38–60. <https://doi.org/10.1086/660686>.
- 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, 173–184. https://doi.org/10.1207/s15374424jccp2802_5.
- Imada, T. (2012). Cultural narratives of individualism and collectivism: A Content analysis of textbook stories in the United States and Japan. *Journal of Cross-Cultural Psychology*, 43, 576–591. <https://doi.org/10.1177/0022022110383312>.
- Jak, S., Oort, F. J., & Dolan, C. V. (2014). Measurement Bias in multilevel data. *Structural Equation Modeling: A Multidisciplinary Journal*, 21, 31–39. <https://doi.org/10.1080/10705511.2014.856694>.
- Jellesma, F. C., Zee, M., & Koomen, H. M. Y. (2015). Children's perceptions of the relationship with the teacher: Associations with appraisals and internalizing problems in middle childhood. *Journal of Applied Developmental Psychology*, 36, 30–38. <https://doi.org/10.1016/j.appdev.2014.09.002>.
- Jia, Y., Ling, G., Chen, X., Ke, X., Way, N., Hughes, D., & Lu, Z. (2009). The influence of student perceptions of school climate on socioemotional and academic adjustment: A comparison of Chinese and American adolescents. *Child Development*, 80, 1514–1530. <https://doi.org/10.1111/j.1467-8624.2009.01348.x>.
- Keller, H. (2013). Attachment and culture. *Journal of Cross-Cultural Psychology*, 44, 175–194. <https://doi.org/10.1177/0022022112472253>.
- Koomen, H. M. Y., & Jellesma, F. C. (2015). Can closeness, conflict, and dependency be used to characterize students' perceptions of the affective relationship with their teacher? Testing a new child measure in middle childhood. *British Journal of Educational Psychology*, 85, 479–497. <https://doi.org/10.1111/bjep.12094>.
- 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, 215–234. <https://doi.org/10.1016/j.jsp.2011.09.001>.
- Lei, H., Cui, Y., & Chiu, M. M. (2016). Affective teacher-student relationships and students' externalizing behavior problems: A meta-analysis. *Frontiers in Psychology*, 7, 1–12. <https://doi.org/10.3389/fpsyg.2016.01311>.
- Lei, H., Cui, Y., & Chiu, M. M. (2018). The relationship between teacher support and students' academic emotions: A meta-analysis. *Frontiers in Psychology*, 8, 1–12. <https://doi.org/10.3389/fpsyg.2017.02288>.
- Little, T. D. (2013). *Methodology in the social sciences. Longitudinal structural equation modeling*. New York, NY, US: Guilford Press.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224–253. <https://doi.org/10.1037/0033-295X.98.2.224>.
- Matsumoto, D., Yoo, S. H., Nakagawa, S., Alexandre, J., Altarriba, J., Anguas-Wong, A. M., ... Yoo, S. H. (2008). Culture, emotion regulation, and adjustment. *Journal of Personality and Social Psychology*, 94, 925–937. <https://doi.org/10.1037/0022-3514.94.6.925>.
- McGrath, K. F., & van Bergen, P. (2015). Who, when, why and to what end? Students at risk of negative student-teacher relationships and their outcomes. *Educational Research Review*, 14, 1–17. <https://doi.org/10.1016/j.edurev.2014.12.001>.
- Mizuta, I., Zahn-Waxler, C., Cole, M., & Hiruma, N. (1996). Cross-cultural study of preschoolers' attachment: Security and sensitivity in Japanese and US dyads. *International Journal of Behavioral Development*, 19, 141–160. <https://doi.org/10.1080/016502596385992>.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, 9, 599–620. https://doi.org/10.1207/S15328007SEM0904_8.
- Oyserman, D., & Lee, S. W. S. (2008). Culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin*, 134, 311–342. doi:<https://doi.org/10.1037/0033-2909.134.2.311>.
- Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, 128, 3–72. <https://doi.org/10.1037/0033-2909.128.1.3>.
- Pianta, R. C. (1992). *New directions for child development: Vol. 57. Beyond the parent: The role of other adults in children's lives*. San Francisco, CA: Jossey Bass.
- Pianta, R. C. (2001). *STRS: Student-teacher Relationship Scale: Professional manual*. Lutz, FL: Psychological Assessment Resources.
- Pianta, R. C., Hamre, B. K., & Stuhlman, M. W. (2003). Relationships between teachers and children. *Handbook of psychology vol. 7: Educational psychology*. Vol. 7. *Handbook of psychology vol. 7: Educational psychology* (pp. 199–234). <https://doi.org/10.1002/0471264385.wei0710>.
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>.
- 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, 346–364. <https://doi.org/10.1177/0143034307078545>.
- Roorda, D. L., Jak, S., Zee, M., Oort, F. J., & Koomen, H. M. Y. (2017). Affective teacher-student relationships and students' engagement and achievement: A meta-analytic update and test of the mediating role of engagement. *School Psychology Review*, 46, 239–261. <https://doi.org/10.17105/SPR-2017-0035.V46-3>.
- 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, 493–529. <https://doi.org/10.3102/0034654311421793>.
- Rothbaum, F., Weisz, J., Pott, M., Miyake, K., & Morelli, G. (2000). Attachment and culture: Security in the United States and Japan. *American Psychologist*, 55, 1093–1104. <https://doi.org/10.1037/0003-066X.55.10.1093>.
- Rubin, K. H., Hemphill, S. A., Chen, X., Hastings, P., Sanson, A., Coco, A. L., & Cui, L. (2006). A cross-cultural study of behavioral inhibition in toddlers: East-West-North-South. *International Journal of Behavioral Development*, 30, 219–226. <https://doi.org/10.1177/0165025406066723>.
- Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, 74, 31–57. <https://doi.org/10.1177/0013164413498257>.
- Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, 75, 243–248. <https://doi.org/10.1007/s11336-009-9135-y>.
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, UK: Sage Publishers.
- Spilt, J. L., Koomen, H. M. Y., Thijs, J. T., & van der Leij, A. (2012). Supporting teachers' relationships with disruptive children: The potential of relationship-focused reflection. *Attachment and Human Development*, 14, 305–318. <https://doi.org/10.1080/14616734.2012.672286>.
- Steenkamp, J. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78–90. <https://doi.org/10.1086/209528>.
- Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology*, 9, 1–12. <https://doi.org/10.1027/1614-2241/a000049>.
- Triandis, H. C. (2001). Individualism-collectivism and personality. *Journal of Personality*, 69, 907–924. <http://doi.wiley.com/10.1111/1467-6494.696169>.
- Triandis, H. C. (2018). *Individualism and collectivism*. New York: Routledge. <https://doi.org/10.4324/9780429499845>.
- Triandis, H. C., Bontempo, R., Villareal, M. J., & Asai, M. (1988). Individualism and collectivism: Cross-cultural perspectives on self/group relationships. *Journal of Personality and Social Psychology*, 54, 323–338. <https://doi.org/10.1037/0022-3514.54.2.323>.
- Tsigilis, N., Gregoriadis, A., & Grammatikopoulos, V. (2018). Research papers in education evaluating the student – Teacher relationship scale in the Greek educational setting: An item parcelling perspective. *Research Papers in Education*, 1522, 1–13. <https://doi.org/10.1080/02671522.2017.1353675>.
- Tsigilis, N., Gregoriadis, A., Grammatikopoulos, V., & Zachopoulou, E. (2018). Applying exploratory structural equation modeling to examine the student-teacher

- relationship scale in a representative Greek sample. *Frontiers in Psychology*, 9, 1–12. <https://doi.org/10.3389/fpsyg.2018.00733>.
- Verschuere, K. (2015). Middle childhood teacher–child relationships: Insights from an attachment perspective and remaining challenges. In G. Bosmans, & K. A. Kerns (Vol. Eds.), *Attachment in middle childhood: Theoretical advances and new directions in an emerging field. New directions for child and adolescent development*. Vol. 148. *Attachment in middle childhood: Theoretical advances and new directions in an emerging field. New directions for child and adolescent development* (pp. 77–91).
- Verschuere, K., & Koomen, H. M. Y. (2012). Teacher–child relationships from an attachment perspective. *Attachment & Human Development*, 14, 205–211. <https://doi.org/10.1080/14616734.2012.672260>.
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 76, 913–934. <https://doi.org/10.1177/0013164413495237>.
- Wu, P., Robinson, C. C., Yang, C., Hart, C. H., Olsen, S. F., Porter, C. L., & Wu, X. (2002). Similarities and differences in mothers' parenting of preschoolers in China and the United States. *International Journal of Behavioral Development*, 26, 481–491. <https://doi.org/10.1080/01650250143000436>.
- Yang, C., Bear, G. G., Chen, F. F., Zhang, W., Blank, J. C., & Huang, X. (2013). Students' perceptions of school climate in the U.S. and China. *School Psychology Quarterly*, 28, 7–24. <https://doi.org/10.1037/spq0000002>.
- Yoon, M., & Millsap, R. E. (2007). Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14, 435–463. <https://doi.org/10.1080/10705510701301677>.
- Zee, M., & de Bree, E. (2017). Students' self-regulation and achievement in basic reading and math skills: The role of student–teacher relationships in middle childhood. *European Journal of Developmental Psychology*, 14, 265–280. <https://doi.org/10.1080/17405629.2016.1196587>.
- Zee, M., de Jong, P. F., & Koomen, H. M. Y. (2017). From externalizing student behavior to student-specific teacher self-efficacy: The role of teacher-perceived conflict and closeness in the student–teacher relationship. *Contemporary Educational Psychology*, 51, 37–50. <https://doi.org/10.1016/j.cedpsych.2017.06.009>.
- Zee, M., & Koomen, H. M. Y. (2017). Similarities and dissimilarities between teachers' and students' relationship views in upper elementary school: The role of personal teacher and student attributes. *Journal of School Psychology*, 64, 43–60. <https://doi.org/10.1016/j.jsp.2017.04.007>.
- Zee, M., Koomen, H. M. Y., & van der Veen, I. (2013). Student-teacher relationship quality and academic adjustment in upper elementary school: The role of student personality. *Journal of School Psychology*, 51, 517–533. <https://doi.org/10.1016/j.jsp.2013.05.003>.
- Zhang, X. (2010). 师幼关系量表的信效度检验[reliability and validity of teacher-child relationship scale]. *Chinese Journal of Clinical Psychology*, 18, 582–583 (doi: 1005-3611(2010)05-0582-02).