Ackermann, Katharina; Martinelli, Anne; Bernhard, Anka; Ueno, Kathrin; Freitag, Christine M.; Büttner, Gerhard; Schmiedek, Florian; Schwenck, Christina

Validation of the network of relationship inventory in female and male adolescents

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
European journal of psychological assessment 36 (2020) 2, S. 220-228

Ackermann, K., Martinelli, A., Bernhard, A., Ueno, K., Freitag, C.M., Büttner, G., Schmiedek, F., & Schwenck, C (2020). Validation of a network of relationship inventory in female and male adolescents. *European Journal of Psychological Assessment, 36(2),* 220-228. https://doi.org/10.1027/1015-5759/a000508

Validation of the Network of Relationship Inventory in Female and Male Adolescents

Katharina Ackermann[1*], Anne Martinelli[1], Anka Bernhard[1], Kathrin Ueno[1], Christine M. Freitag[1], Gerhard Büttner [2], Florian Schmiedek[3], & Christina Schwenck[1&4]

Authors Note

[1]Department of Child and Adolescence Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt, Goethe University, Frankfurt am Main

[2]Department of Psychology, Goethe University, Frankfurt am Main

[3]German Institute of International Educational Research (DIPF), Frankfurt am Main

[4]Department of Psychology, Justus-Liebig-University, Gießen

Conflict of interest

All authors declare no competing interests.

Role of funding source and acknowledgement

This research was supported by the European Union (FP7 grant no. 602407; FemNAT-CD). We are grateful to all research assistants for their help in the data collection and to all schools, participants and their families for their friendly collaboration and support for this research.

*Corresponding author: Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Frankfurt, Goethe University, Deutschordenstraße 50, 60528 Frankfurt am Main, Germany. Phone number: +496963017178. Fax number: +496963015843. E-mail address: katharina.ackermann@kgu.de

Article type: Original article

Word count: 5705 words

Tables:

Table 1. Results on confirmatory factor analysis and measurement invariance testing
Table 2. Gender differences in factor means among the nine factors

Figure:

Figure 1. The final nine-factor model (ModelC) with completely standardized item factor loadings and correlations (only shown when *r.* > .30) among factors for females/males. Fixed items (rectangles) depicted by broken lines; Latent factors (circles) are: COM=companionship, INT=intimate disclosure/support, SAT=satisfaction, APP=approval, PRE=pressure, CON=conflict, CRI=criticism, DOM = dominance, EXC = exclusion.

Abstract

Friendships and their different qualities have been shown to be important for adolescents' socio-emotional development and psychological adjustment. In empirical research on such friendship qualities, the Network of Relationship Inventory (NRI-RQV) is a widely used questionnaire. Here, we conduct an extensive validation of a German version of the NRI-RQV, investigating its factor structure, reliability, and concurrent validity, in a sample of *N*=679 adolescents aged 13 to 18 years. Applying multi-group confirmatory factor analysis, we further test whether the factor structure of the friendship quality construct holds across groups of males and females. Results showed that a structure with nine correlated first-order factors fit the data well, indicating nine distinct friendship qualities in males and females. Measurement invariance testing suggested the same underlying friendship quality construct, albeit differences in mean scores per gender. As evidence for concurrent validity closeness and discordant friendship qualities showed expected correlations with empathy and social problems, respectively, but not with aggressive behavior. Overall, results indicate good psychometric properties for the German version of the NRI-RQV as a measure of friendship qualities in both, males and females.

Keywords: friendship quality questionnaire, multi-group confirmatory factor analysis, invariance testing, gender differences

Validation of the Network of Relationship Inventory in Female and Male Adolescents

During late childhood and adolescence, the influence of peer relationships on social development becomes increasingly important. Adolescents spend increased time with their peers, due to a change in the nature of friendships - from playmates in childhood to confidants who provide emotional closeness in adolescence (Lansford, Criss, & Pettit, 2003). Both theory and empirical findings highlight the importance not only of quantitative aspects of friendships in adolescence, such as the number of friends, but most importantly friendship quality (Bagwell, 2005; Berndt, 2004; Bukowski, Hoza, & Boivin, 1994). Friendship quality refers to distinctive positive and negative features in close friendships, which both can exist equally in one friendship (Berndt, 2004; Furman & Buhrmester, 1985). Developmental and clinical psychological studies focus on how positive and negative friendship quality is related to individual characteristics and behaviors, and how it influences therapeutic processes as well as later wellbeing (e.g., Baker & Hudson, 2014; Mundt & Zakletskaia, 2014). The assessment of distinct friendship qualities is therefore highly relevant for adolescent development.

The most common form of friendship quality assessment is through self-report questionnaires (e.g., Waldrip & Malcolm, 2008). One of the most frequently used self-report questionnaires to assess positive and negative friendship quality (e.g., Chow, Ruhl, & Buhrmester, 2013; Nieder & Seiffge-Krenke, 2001; van Aken & Asendorpf, 1997) is the Network of Relationship Inventory – Relationship Quality Version (NRI-RQV) by (Buhrmester & Furman, 2008). The NRI-RQV was developed to broaden the assessment of negative friendship quality features, as other friendship quality questionnaires suggested multiple-factor solutions with only one scale (conflict) representing negative and four or more scales representing positive friendship qualities (e.g., *Friendship Quality Scale*, Bukowski et

al., 1994; *Friendship Quality Questionnaire*, Parker & Asher, 1993; *Friendship Quality Measure*, Grotpeter & Crick, 1996). However, a questionnaire, which equally represents positive and negative friendship qualities, has not been extensively tested yet, although some items of the NRI are already in use in German populations (e.g., Nieder & Seiffge-Krenke, 2001; van Aken & Asendorpf, 1997) and distinct negative qualities seem to be especially relevant when interested in associations with well-being (e.g., Parker, Rubin, Erath, Wojslawowicz, & Buskirk, 2005). Buhrmester and Furman (2008) describe the items and subscales of the original NRI-RQV in an unpublished study of N=223 11- to 12-year-old children. The NRI-RQV includes 30 items, rated on a 5-point Likert scale from "1=never to hardly at all" to "5=always or extremely much" and describes behavior that occurs within the context of the relationship (e.g., "How often do you depend on your friend for help, advice, or sympathy?"). The positive friendship quality scale "closeness" is composed of companionship, intimate disclosure, satisfaction, emotional support, and approval. The five subscales pressure, conflict, criticism, dominance, and exclusion can be subsumed under a negative friendship quality scale "discord". The five positive and negative qualities showed an internal consistency of α=.68 to α=.95, and α=.65 to α=.90, respectively (Buhrmester & Furman, 2008). Supporting a two factor solution, the *NRI - Behavioral Systems Version*, a version related to the NRI-RQV and assessing conceptualizations of close relationships, shows a second-order factor structure with eight subscales loading on two higher-order factors "support" and "negative interaction" (Furman & Buhrmester, 2009). Kouwenberg, Rieffe, and Banerjee (2013) also found a two-factor solution using principle component analysis in their *Best Friend Index* with one "positive friendship factor" and one "negative friendship factor".

Results on gender differences in the assessment of friendship quality are inconsistent. Studies showed that females scored significantly higher in positive qualities and significantly lower in negative friendship qualities than males (e.g., Chow et al., 2013). This difference was

also found in a longitudinal study by (De Goede, Branje, & Meeus, 2009). Here, gender differences persisted, and females focused more on self-disclosure and empathy whereas male friendship was based on companionship, competition, and control. In contrast, gender differences might vary depending on the applied questionnaire or studied sample. In a review on gender differences, none were found in studies with smaller sample sizes (Rudolph & Rose, 2006). Additionally, some studies report gender differences in positive, but not in negative, friendship qualities (Bukowski et al., 1994; Kouwenberg et al., 2013; Parker & Asher, 1993). Despite multiple results on gender differences, no study has tested for measurement invariance in the friendship quality construct. However, comparisons across gender may be invalid if a factor being measured is not invariant across females and males.

Related constructs to closeness and discordant qualities such as socio-emotional functioning (empathy and social problems) and psychopathological symptoms (aggressive behavior) have been used to evaluate concurrent validity (e.g., Kouwenberg et al., 2013). Adolescents who show more empathic abilities maintain friendships characterized by more care, companionship, validation, and fewer conflicts (Chow et al., 2013; Smith & Rose, 2011). In contrast, aggressive behavior is associated with more conflict and less closeness within friendships (Bagwell & Coie, 2004), although this association has been shown differently according to various forms of aggression (e.g., Rose, Swenson, & Carlson, 2004a). In addition, having social problems, such as being socially withdrawn from or bullied by others, is associated with discordant friendship qualities (Rubin, Wojslawowicz, Rose-Krasnor, Booth-LaForce, & Burgess, 2006).

Despite the high importance of friendship quality with respect to research and practice (e.g., Véronneau, Trempe, & Paiva, 2014), to our knowledge, an extensive validation and confirmation of the NRI-RQV, and a validation and confirmation of a German friendship quality questionnaire in particular, is lacking. Little is known about whether the mostly used

two-factor structure (including both positive and negative features) in friendship quality questionnaires reflects the construct of friendship quality equally well for females and males. Therefore, the aim of the current study was to analyze the factor structure, reliability, and concurrent validity of the German version of the NRI-RQV, including tests of measurement invariance across gender. We hypothesized the same factor structure as found in the original version with ten subscales loading on two higher-order factors. In line with previous results, we hypothesized that females would score higher in closeness and lower in discordant friendship qualities than males. Finally, with respect to concurrent validity, we expected positive correlations between the closeness scale and empathy, as well as positive correlations between the discord scale and social problems and aggressive behavior.

Method

Participants and Procedures

The sample comprised *N*=679 (*N*=374 females) participants between 13 and 18 years (*M*=14.63 years, *SD*=1.38) of two independently run studies. In the first study *N*=598 adolescents were recruited from 17 public schools in three different federal states in Germany. Paper-pencil questionnaires were filled out during class. Parents filled out questionnaires concerning parental educational status. A second sample (*N*=81) was recruited within an on-going European wide case-control study on conduct disorder (FemNAT-CD). Only control participants, who did not have any current psychiatric diagnoses, were included in the current analysis. The ethics committee of the University Hospital Frankfurt approved both studies (file no.: 438/13 and 445/13). Written informed consent was obtained from all participants and their parents. Of the total sample, *N*=430 parents completed questionnaires on socio-demographic data including country of birth, educational status, and income [see Electronic Supplementary Material (ESM), Table1].

Measures

Friendship quality. The original *NRI-RQV* (Buhrmester & Furman, 2008) was forward-backward translated into/from German to ensure the semantic equivalence of the German and the English versions. The process involved two independent researchers: one native German speaker who was fluent in English translated the original version into German. The second researcher, a native English speaker who was fluent in German, translated this German version back into English. Both researchers discussed resulting differences between the versions. Discrepancies were solved in alignment with the original English meaning. The participants were instructed not to include siblings or romantic partners as best friends.

Concurrent validity measures. Within the same test-time point, the following measures were collected to assess concurrent validity. In line with previous literature two subscales of the *Interpersonal Reactivity Index* (IRI) by Davis (1980) was administered to assess cognitive (*perspective taking; PT*) and affective *(empathic concern; EC)* empathy (Chow et al., 2013). The 28-item self-rating questionnaire is answered on a 5-point Likert-scale (0=”does not describe me well” to 4=“describes me very well”) with an internal consistency of PT α=.53 and EC α=.38 (N=675) in the current sample.

Social problems and aggressive symptoms were assessed using the “social problem” and “aggressive behavior” subscales of the *Youth Self Report* (YSR; Achenbach, 1991; Arbeitsgruppe Deutsche Child Behavior Checklist, 1998). The YSR is a 113 item self-report questionnaire for children and adolescents between 11 and 18 years assessing a wide spectrum of psychiatric symptoms rated on a 3-point Likert-scale. In the current sample the “social problem” and the “aggressive behavior” subscales showed an internal consistency of α=.68 and α=.87, respectively (N=589).

Data Analysis

Confirmatory factor analysis (CFA). This study uses CFA in the full sample of females and males to replicate the original second-order model proposed by the original authors, with ten subscales loading on two higher-order factors "closeness" and "discord" (Model A). Further CFAs were planned if the factor structure could not be reproduced with an acceptable fit. The CFA was performed with robust maximum likelihood (MLR) estimation to address non-normally distributed data, with marked skewedness and kurtosis, and to avoid potential bias of the full information maximum likelihood (FIML) estimation method (Brown, 2015).

Model goodness of fit was evaluated using several fit indices. The Comparative Fit Index (CFI) and Tucker Lewis fit index (TLI) values are acceptable when > .90 and good when > .95. Further, an acceptable fit was given when the Root Mean Square Error of Approximation (RMSEA) was below .05 and the Standardized Root Mean Square Residual (SRMSR) less than .08 (Marsh, Hau, & Wen, 2004). Satorra-Bentler (S-B) χ^2 tests were conducted to test whether differences in model fit were significant (Bryant & Satorra, 2012) and when likelihood ratio tests were used for testing hypotheses. In addition, the above mentioned goodness of fit indices were also considered, because such likelihood ratio tests are sensitive to large sample sizes and can become significant with only small impairments of model fit.

Scale reliability. Scale reliability was estimated for males and females separately, for the best fitting model, by calculating test-retest reliability and internal consistency. The participants were recruited from Study 1 and asked to fill out the questionnaire again after one year. Test-retest reliability was considered good when correlation between test points was r > .70. Internal consistency was considered good with McDonald's coefficient Omega > .70.

Testing measurement invariance. After conducting the CFA in females and males separately (Model 0a and 0b), a multi-group CFA, subsequently constraining the model from the least strict to the strictest model, was chosen to evaluate measurement invariance (Brown, 2015). Test of measurement invariance can be conducted to test whether the factor structure is invariant across groups, allowing a meaningful comparison of females' and males' friendship quality mean scores. First, the least strict model tests for *configural invariance (Model 1)*, that is, whether the same configuration of items load on the same factors in females and males, while allowing parameter values (e.g., factor loadings and intercepts) to differ between groups. Second, *metric invariance (Model 2)* requires that (unstandardized) factor loadings are equal across groups. If there is no significant loss of model fit between *Model 1* and *Model 2*, it can be assumed that in both groups the latent factors have a sufficiently similar substantive interpretation. Third, *scalar invariance (Model 3)* requires that the intercepts of the latent variables are equal across groups. If scalar invariance holds, the means of the latent factors can be compared between groups. Forth, error variances were constrained to be equal in both groups (*strict invariance*, *Model 4)*. If this model does not fit the data well, it can be assumed that reliability might differ between groups. To test whether each subsequently stricter model still fit the data, S-B χ^2 test (to test whether differences in model fit were significant) and the above mentioned goodness of fit indices (CFI, TLI, RMSEA, SRMSR) were considered (Bryant & Satorra, 2012; Marsh et al., 2004).

Concurrent validity. To test concurrent validity each latent factor of the NRI-RQV was correlated with the total scores of cognitive and affective empathy, social problems, and aggression. The best fitting multi-group CFA model was used to correlate the friendship quality factors with empathy, social problems and aggressive behavior subscales.

For all statistical analyses, Mplus Version7 was used (Muthén & Muthen, 2015).

Results

Confirmatory Factor Analysis

Results of the CFA on the whole sample (N=679) did not show an acceptable fit for the proposed higher-order factor model with ten latent first-order factors (Model A; see Table 1). Therefore, the first-order structure was tested without the second-order factor structure. This model showed an acceptable and significantly better fit to the data than the model with two higher-order factors (Model B).

Due to a high correlation between two latent factors (intimate disclosure and support; r = .99), these two factors were combined. The resulting model with nine factors fit the data acceptably well (Model C). The second-order factor structure was then tested again for the remaining nine subscales, resulting in acceptable model fit (Model D). Due to an acceptable model fit of Model C and D, both models were compared in each step of measurement invariance testing using a S-B χ^2 test and the above-mentioned goodness of fit indices (CFI, TLI, RMSEA, SRMSR). In all steps, Model C resulted in a significantly better model fit, and numerically better values of the fit indices, than Model D (see ESM, Table 2). Therefore, Model C (Figure 1) was chosen for further analysis.

Scale Reliability

ESM Table 3 shows the items of the NRI, item factor loadings, item difficulty, and coefficient Omega for internal consistency for Model C. Omegas ranged from .83 to .94 for females and from .77 to .92 for males, indicating good internal consistency for all factors. All items loaded on the expected factors with item factor loadings ranging between .38 and .90 for females and .30 and .83 for males. Item difficulty ranged from .65 to .88 for females and .51 to .84 for males in companionship, intimate disclosure/support, satisfaction and approval and from .08 to .47 for females and .17 to .48 for males in pressure, conflict, criticism,

dominance, and exclusion. Test-retest reliability was estimated in a sub-sample of *N*=77 (46 females; age 13-18, *M*=14.25, *SD*=1.36). No factor correlated higher than .70 between T1 and T2, although most factors for females and some for males reached significance (see ESM Table 4).

Measurement Invariance Testing

Before testing measurement invariance with a multiple-group analysis, Model C was tested separately in each group with a single CFA. Table 1 presents the fit indices and S-B χ^2 -difference test results. The single CFA model fit the data acceptably for both females (Model 0a, *N*=374) and males (Model 0b, *N*=305). In both groups, all freely estimated factor loadings were statistically significant (all *p*s < .001; completely standardized factor loadings ranged from .38 and .84 in males and .43 to .89 in females). Results of the multi-group CFA with configural invariance (Model 1) showed an acceptable fit, indicating a comparable factor structure between groups. In addition, the restriction of equal factor loadings (Model 2) was accompanied by a non-significant change of fit, confirming the assumption of metric invariance. After setting latent variable intercepts equal (Model 3), the S-B χ^2 -test reached significance. When the error indicators were set to equal (Model 4), though, the model showed a statistically significant decrease in overall model fit as well as poor fit in the other indices. Therefore, Model 3 was still considered a tenable assumption, implying that latent factor means and correlations (see ESM, Table 5) can be compared across groups in a meaningful way. As shown in Table 2, females scored significantly higher on companionship, intimate disclosure/support, satisfaction, and approval than males and males scored significantly higher on pressure, conflict, criticism, and dominance than females. Although the difference of the exclusion scale was in the presumed direction, this difference did not reach significance.

244

245 **Concurrent Validity**

246 Descriptive statistics on all concurrent validity measures, their correlations between
247 the latent factors of the NRI-RQV, as well as summarized results on analyses relating to
248 concurrent validity can be found in the ESM Table 6 and 7, respectively. With respect to
249 empathy, significant positive correlations emerged between males' cognitive empathy and
250 intimate disclosure/support and approval, while significant negative correlations between
251 cognitive empathy and criticism were found. With respect to affective empathy, small positive
252 correlations were found between companionship and satisfaction for females and intimate
253 disclosure/support for males. Results show no correlations between discord friendship
254 qualities and affective empathy, with the exception of a small negative correlation between
255 exclusion and affective empathy in females. Social problems correlated negatively with
256 closeness and positively with discord friendship qualities for females and negatively with
257 intimate disclosure/support and positively with exclusion for males. Except for males
258 reporting significantly less conflict (r= - .29, p < .05), and for females reporting significantly
259 less exclusion (r= - .26, p < .05), when showing more aggressive behavior, no significant
260 correlations were found with aggressive behavior.

261

262 **Discussion**

263 Overall, the results of the current study show that the German NRI-RQV is a valid and
264 reliable measure to assess friendship quality in males and females. Using a confirmatory
265 factor analysis, the best fitting model for the current sample includes nine distinct and
266 partially correlated factors: companionship, intimidate disclosure/support, satisfaction,
267 approval, pressure, conflict, criticism, dominance, and exclusion. These qualities are in line
268 with the original construct, except that the factors intimate disclosure and support were

combined, due to a high correlation with each other. This high correlation may be explained by similar goals when "telling a best friend everything that one is going through" (intimate disclosure) and "asking the best friend for support with personal problems" (emotional support). The higher-order factor structure could be replicated but was not shown to be the best fitting model. A more distinct concept of friendship quality with nine first-order factors could be explained by sample characteristics. The present study used a larger age range and older adolescents (13-18 years) in comparison to the original study (11-12 years). Friendship qualities change and develop during adolescents' development. The need for intimacy, self-disclosure, and support by the best friend becomes increasingly important with age whereas companionship is already highly important in childhood (Parker et al., 2005; Rudolph & Rose, 2006). Further, especially in female friendships, the experience of dominance becomes less present, whereas in male friendships conflict and pressure remain relatively stable throughout adolescence (De Goede et al., 2009). With respect to reliability, results indicate good internal consistency, similar factor loadings and item difficulties in both genders. This indicates that all items "function" the same way for females and males. In contrast, a higher test-retest reliability for females compared to males might imply that males' are not as stable as females' friendship qualities. However, the poor test-retest reliability should be interpreted with caution, due to the small sample size.

Although friendship quality questionnaires have been equally used for males and females, and despite common knowledge about mean differences in friendship quality magnitudes between genders, to our knowledge, this was the first study to assess gender invariance in the assessment of friendship quality. Results support the hypothesis and previous literature that the core provisions of friendship are similar in males and females despite different magnitudes in friendship characteristics (e.g., Carlson Jones, 1991). As

concluded by other authors before, it is possible that the phrasing of positive friendship qualities within questionnaires captures a more feminine understanding of friendships (Duck & Wright, 1993). This could mean that males might feel as close as females to their friends, but understand and show this closeness differently (Parker et al., 2005). Higher mean scores on discordant qualities among males lead to the question of how and why males see their friendships in a more discordant way. Rudolph and Rose (2006) point out that males are more competitive due to a larger friendship group in which they strive to protect or/and gain social standing. In contrast, females focus more on one best friend whose friendship they protect with more self-disclosure, more polite discussions and compromising (Rudolph & Rose, 2006). This interpretation appears contradictory when considering the results on empathy. The relationship between perspective taking and more intimate disclosure, support, and approval in friendships, as well as less perceived criticism, was especially found in males. Therefore, a second reason for the gender differences in discordant qualities could be that males are more willing to express conflict and dominance in self-report questionnaires compared to females, because males do not see those features as something "negative" within a friendship. Indeed, dominant behavior might even be something they see as socially desired.

Results of the current study only partially confirmed relations with other constructs. Supporting concurrent validity of the friendship quality construct, social problems correlated negatively with closeness and positively with most discordant friendship qualities. This supports the assumption that not being in contact with peers due to social withdrawal or victimization is also related to higher discord and lower closeness qualities in best friendships (Kendrick, Jutengren, & Stattin, 2012; Rubin et al., 2006). Additionally, results on empathy suggest that those with higher cognitive empathy maintain more intimacy in friendships (Chow et al., 2013; Smith & Rose, 2011), which seems to be especially true for males.

However, although in the expected direction, only a few small correlations between affective empathy and friendship qualities emerged. In contrast to past research, this study differentiated between cognitive and affective empathy, whereas past research used one total score including cognitive and affective empathic abilities (Chow et al., 2013). Against expectation, aggressive symptoms were rather independent of the perceived friendship quality construct, which may indicate poor construct validity. However, the composition of this population-based sample may have resulted in low correlations with aggression due to a low variance in occurrence of aggressive behavior. In addition, it has been reported that relational and overt aggression are distinctly related to discord and closeness friendship qualities (Ackermann et al., 2018; Deptula & Cohen, 2004; Kamper & Ostrov, 2013; Rose, Swenson, & Waller, 2004b). Future studies could therefore include samples with a larger range and different forms of aggressive behavior.

Limitations and Future Directions

Despite, major advantages of this study including a large sample size and a detailed elaboration of a German translation of a widely used friendship quality questionnaire; a few issues should be considered.

First, results may differ between age groups. Although De Goede et al. (2009) did not find a decline in negative interactions in friendships, friendships in males and females did become more intimate and supportive with age. Especially in late adolescence, there might be a shift in importance from friendships to intimate romantic relationships. Future research with larger sample sizes for each age should address the effect of age. Second, the relatively poor test-retest reliability could be due to the long duration between test-points. It suggests that a single year in an adolescents' life leads to multiple developmental changes, also with respect to relationship quality. Therefore, in future studies a shorter timeframe for the test-retest should

be considered. Third, the empathic concern and perspective taking subscale of the IRI show poor internal consistency. However, the IRI is a widely used self-report measure on empathy and was chosen to allow comparison of the present study results to previous findings. Forth, future studies focusing more strongly on concurrent validity may consider using SEM and latent factors for all constructs, instead of correlations scale scores. Lastly, information on educational background and economical status was present only for a subsample (those whose parents took part in the study), which leads to limited knowledge of the representativeness of this sample. However, due to the combination of two samples, overall sample size was increased, which improved statistical analyses.

Despite these limitations, the current study extends previous literature by showing that the German NRI-RQV is a useful questionnaire assessing friendship quality. This study was the first to extensively assess the factor structure and factorial measurement invariance between genders, indicating that the mean factor differences between males and females are not based on a different underlying friendship quality construct. Therefore, the German NRI-RQV may be especially useful for future studies on gender differences in associations between friendship qualities, socio-emotional functioning, and psychopathology.

Literature

Achenbach, T. M. (1991). *Manual for the Youth Self-Report and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Ackermann, K., Büttner, G., Bernhard, A., Martinelli, A., Freitag, C. M., & Schwenck, C. (2018). Freundschaftsqualitäten und unterschiedliche Formen aggressiven Verhaltens bei Jungen und Mädchen im späten Kindes- und Jugendalter. *Kindheit Und Entwicklung*, *27*(2), 81–90.

Arbeitsgruppe Deutsche Child Behavior Checklist. (1998). *Fragebogen für Jugendliche; deutsche Bearbeitung der Youth Self-Report Form der Child Behavior Checklist (YSR). Einführung und Anleitung zur Handauswertung [Manual for the child behavior checklist - youth self report form and profile].* Köln.

Bagwell, C. L. (2005). Friendship quality and perceived relationship changes predict psychosocial adjustment in early adulthood. *Journal of Social and Personal Relationships*, *22*(2), 235–254.

Bagwell, C. L., & Coie, J. D. (2004). The best friendships of aggressive boys: Relationship quality, conflict management, and rule-breaking behavior. *Journal of Experimental Child Psychology*, *88*(1), 5–24.

Baker, J. R., & Hudson, J. L. (2014). Friendship quality and social information processing in clinically anxious children. *Child Psychiatry & Human Development*, *45*(1), 12–23.

Berndt, T. J. (2004). Friendship and three A's (aggression, adjustment, and attachment). *Journal of Experimental Child Psychology*, *88*(1), 1–4.

Brown, T. A. (2015). Confirmatory factor analysis for applied research (2nd ed.). New York.

Bryant, F. B., & Satorra, A. (2012). Principles and practice of scaled difference chi-square testing. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(3), 372–398.

Buhrmester, D., & Furman, W. (2008). *The Network of Relationship Inventory: Relationship Quality Version*. University of Texas.

Bukowski, W. M., Hoza, B., & Boivin, M. (1994). Measuring friendship quality during pre- and early adolescence: The development and psychometric properties of the Friendship Qualities Scale. *Journal of Social and Personal Relationships*, *11*(3), 471–484.

Carlson Jones, D. (1991). Friendship satisfaction and gender: An examination of sex differences in contributors to friendship satisfaction. *Journal of Social and Personal Relationships*, *8*(2), 167–185.

Chow, C. M., Ruhl, H., & Buhrmester, D. (2013). The mediating role of interpersonal competence between adolescents' empathy and friendship quality: A dyadic approach. *Journal of Adolescence*, *36*(1), 191–200.

Davis, M. H. (1980). *A multidimensional approach to individual differences in empathy*. Austin, Texas: JSAS Catalog of Selected Documents in Psychology.

De Goede, I. H. A., Branje, S. J. T., & Meeus, W. H. J. (2009). Developmental changes and gender differences in adolescents' perceptions of friendships. *Journal of Adolescence*, *32*(5), 1105–1123.

Deptula, D. P., & Cohen, R. (2004). Aggressive, rejected, and delinquent children and adolescents: A comparison of their friendships. *Aggression and Violent Behavior*, *9*(1), 75–104.

Duck, S., & Wright, P. H. (1993). Reexamining gender differences in same-gender friendships: A close look at two kinds of data. *Sex Roles*, *28*(11/12), 709–727.

Furman, W., & Buhrmester, D. (1985). Children's perceptions of the personal relationships in their social networks. *Developmental Psychology*, *21*(6), 1016–1024.

Furman, W., & Buhrmester, D. (2009). The Network of Relationships Inventory: Behavioral Systems Version. *International Journal of Behavioral Development*, *33*(5), 470–478.

Grotpeter, J. K. & Crick, N. R. (1996). Relational aggression, overt aggression, and friendship. *Child Development, 67*(5), 2328 - 2338.

Kamper, K. E., & Ostrov, J. M. (2013). Relational aggression in middle childhood predicting

adolescent social-psychological adjustment: The role of friendship quality. *Journal of Clinical Child and Adolescent Psychology*, *42*(6), 855–862.

Kendrick, K., Jutengren, G., & Stattin, H. (2012). The protective role of supportive friends against bullying perpetration and victimization. *Journal of Adolescence*, *35*(4), 1069–1080.

Kouwenberg, M., Rieffe, C., & Banerjee, R. (2013). Developmetrics a balanced and short Best Friend Index for children and young adolescents. *European Journal of Developmental Psychology*, *10*(5), 634–641.

Lansford, J. E., Criss, M. M., & Pettit, G. S. (2003). Friendship quality, peer group affiliation, and peer antisocial behavior as moderators of the link between negative parenting and adolescent externalizing behavior. *Journal of Research on Adolescence*, *13*(2), 161–184.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341.

Mundt, M. P., & Zakletskaia, L. I. (2014). That's what friends are for: Adolescent peer social status, health-related quality of life and healthcare costs. *Applied Health Economics and Health Policy*, *12*(2), 191–201.

Muthén, B. O., & Muthen, L. (2015). M*plus* User's Guide (7 ed., pp. 1–876). Los Angeles.

Nieder, T., & Seiffge-Krenke, I. (2001). Psychosoziale Determination depressiver Symptome im Jugendalter: Ein Vergleich der Geschlechter [Psychosocial determination of depressive symptoms in adolescence: A gender comparison]. *Praxis Der Kinderpsychologie Und Kinderpsychiatrie*, *50*(5), 342–359.

Parker, J. G., & Asher, S. (1993). Friendship and friendship quality in middle childhood: Links with peer group acceptance and feelings of loneliness and social dissatisfaction. *Developmental Psychology*, *29*(4), 611–621.

Parker, J. G., Rubin, K. H., Erath, S. A., Wojslawowicz, J. C., & Buskirk, A. A. (2005). Peer relationships, child development, and adjustment: A developmental psychopathology perspective. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychology theory and method* (2nd ed., Vol. 1, pp. 419–493). Hoboken, NJ.

Rose, A. J., Swenson, L. P., & Carlson, W. (2004a). Friendships of aggressive youth: Considering the influences of being disliked and of being perceived as popular. *Journal of Experimental Child Psychology*, *88*(1), 25–45.

Rose, A. J., Swenson, L. P., & Waller, E. M. (2004b). Overt and relational aggression and perceived popularity: Developmental differences in concurrent and prospective relations. *Developmental Psychology*, *40*(3), 378–387.

Rubin, K. H., Wojslawowicz, J. C., Rose-Krasnor, L., Booth-LaForce, C., & Burgess, K. B. (2006). The best friendships of shy/withdrawn children: Prevalence, stability, and relationship quality. *Journal of Abnormal Child Psychology*, *34*(2), 139–153.

Rudolph, K. D., & Rose, A. J. (2006). A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin*, *132*(1), 98–131.

Smith, R. L., & Rose, A. J. (2011). The “cost of caring” in youths' friendships: Considering associations among social perspective taking, co-rumination, and empathetic distress. *Developmental Psychology*, *47*(6), 1792–1803.

van Aken, M. A. G., & Asendorpf, J. B. (1997). Support by parents, classmates, friends and siblings in preadolescence: Covariation and compensation across relationships. *Journal of Social and Personal Relationships*, *14*(1), 79–93.

Véronneau, M.-H., Trempe, S.-C., & Paiva, A. O. (2014). Risk and protection factors in the peer context: How do other children contribute to the psychosocial adjustment of the adolescent? *Ciência & Saúde Coletiva*, *19*(3), 695–705.

Waldrip, A. M., & Malcolm, K. T. (2008). With a little help from your friends: The

importance of high-quality friendships on early adolescent adjustment. *Social Development*, *17*(4), 832–852.

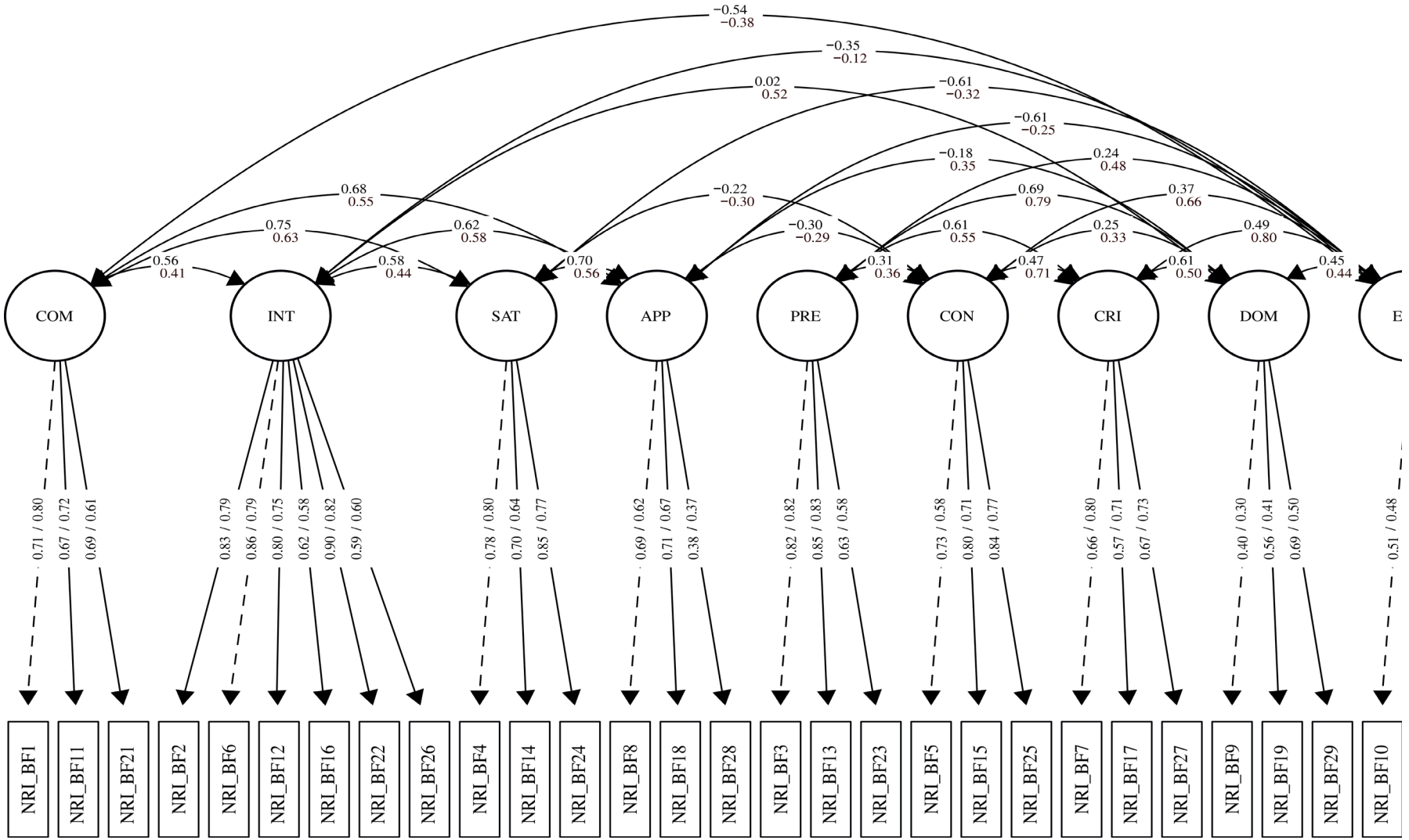

Figure 1. The final nine-factor model (ModelC) with completely standardized item factor loadings and correlations (only shown when *r.* > .30) among factors for females/males. Fixed items (rectangles) depicted by broken lines; Latent factors (circles) are: COM=companionship, INT=intimate disclosure/support, SAT=satisfaction, APP=approval, PRE=pressure, CON=conflict, CRI=criticism, DOM = dominance, EXC = exclusion.

Table 1. Results on confirmatory factor analysis and measurement invariance testing

Model		χ^2	*df*	RMSEA [90% CI]	CFI	TLI	SRMR	TRd/*Δdf*	*p*
Confirmatory factor analysis									
A	Ten 1st order factors on two 2nd order factors	1067.69	394	.05 [.05 - .05]	.89	.88	.08	-	-
B	Ten 1st order factors	737.21	360	.04 [.03 - .04]	.94	.92	.05	311 (34)	.000
C	Nine 1st order factors	769.65	369	.04 [.04 - .04]	.93	.92	.05	31 (9)	.000
D	Nine 1st order factors on two 2nd order factors	945.75	395	.04 [.04 - .05]	.91	.90	.07	177 (26)	.000
Measurement invariance testing of Model C									
0a:	Females	599.49	369	.04 [.03 - .05]	.94	.92	.05	-	-
0b:	Males	613.44	369	.05 [.04 - .05]	.90	.88	.06	-	-
1:	Unconstrained	1212.99	738	.04 [.04 - .05]	.92	. 90	.06	-	-
2:	Factor loadings equal	1239.16	759	.04 [.04 - .05]	.92	.91	.06	28.37 (21)	.130
3:	Factor loadings and intercepts equal	1299.96	780	.04 [.04 - .05]	.91	.90	.06	63.25 (21)	.000
4:	Factor loadings, intercepts, error variance equal	1446.50	810	.05 [.04 - .05]	.89	.89	.07	101.54 (30)	.000

Notes. χ^2 =Chi-Square, *df*=degrees of freedom, RMSEA=Root Mean Square Error of Approximation, CFI=Comparative Fit Index, TLI=Tucker-Lewis Index, SRMR=Root Mean Square Error of Approximation, TRd/*Δdf* =Santorra-Bentler Chi Square/ difference in degrees of freedom, *p*=value of significance **[Author: Please check whether *p*= .000 should be *p*<.001?]**.

Table 2. Gender differences in factor means among the nine factors

		com	int_sup	sat	app	pre	con	cri	dom	exc
M	Female	4.03	3.95	4.49	3.70	1.75	1.87	1.44	2.85	1.57
	Male	3.89	3.04	4.37	3.34	1.98	1.96	1.73	2.96	1.66
	p	.023	.000	.013	.000	.000	.043	.000	.000	.520
	d	- 0.22	- 1.07	- 0.20	- 0.59	0.29	0.17	0.45	0.39	.051

Notes. com=companionship, int_sup=intimate disclosure/support, sat=satisfaction, app=approval, pre=pressure, con=conflict, cri=criticism, dom=dominance, exc=exclusion. **[Author: Please check whether *p*= .000 should be *p*<.001?]**.