Article

Parent and Teacher Assessments of Social-Emotional Competence in Three-Year-Old Children: Does Sibling Status Matter?

Carina Schönmoser[1] ⓘ**, Claudia Karwath**[1] ⓘ**, and Timo Gnambs**[1] ⓘ

Abstract

Valid information on early social-emotional competence is essential to diagnose, treat, and prevent behavioral problems in children and adolescents. Particularly in young children, social-emotional competence is frequently measured using parent and teacher ratings that frequently exhibit low agreement. Therefore, the present study on n = 532 three-year-olds (47% girls) examined whether sibling status might explain discrepancies between the two informant groups. First, multi-trait multi-informant analyses explored the construct validity of a short measure of three facets of social-emotional competence. Then, group comparisons evaluated the size of the observed method effects for only children and children with siblings. Results showed low convergent validity between parent and teacher ratings for aggressive behavior, cooperative behavior, and emotional self-regulation. Sibling status in the family contributed little to the observed discrepancies between parents and teachers. Thus, a comprehensive assessment of social-emotional competence in children requires a multi-informant approach to capture the construct breadth.

Social-emotional competence consists of various skills, knowledge, and abilities facilitating socially competent behavior (Kanning, 2002). The foundations of social-emotional competence are already laid in early childhood. Developmental problems or delays as early as age three can

[1]Leibniz Institute for Educational Trajectories, Bamberg, Germany

Corresponding Author:
Carina Schönmoser, Leibniz Institute for Educational Trajectories, Wilhelmsplatz 3, Bamberg 96047, Germany,
Email: carina.schoenmoser@lifbi.de

lead to problems up to adolescence and even adulthood. For example, high social-emotional competence throughout the life course has been repeatedly shown to predict social and academic success (e.g., Barry & Wigfield, 2002; Denham et al., 2014; 2009; Greco & Morris, 2005; Stepp et al., 2011). Therefore, it is crucial to implement intervention strategies at an early stage and, thus, to validly assess social-emotional competence as early as kindergarten. Researchers typically rely on observer ratings from different informants (multi-informant perspective) such as kindergarten teachers and parents to obtain an overall impression of a child's social-emotional competence. However, parent and teacher ratings of children's social-emotional competence often correlate rather poorly (Achenbach et al., 1987; Rescorla et al., 2012). Low correlations could be problematic if decisions about early interventions, such as delaying school entry, behavioral therapy, or psychotherapy, are made based on a single informant assessment. Reasons for these low correlations might be, for example, different social settings that different informants observe or informant bias. As of yet, no study has considered sibling status as a source of informant bias, although stereotypes for children with and without siblings, such as only children being less socially competent, exist (Thompson, 1974). When considering the ratings of social-emotional competence and sibling status in correlation analyses, it is not clear whether the different ratings are due to a child's social-emotional development being affected by growing up with or without a sibling or whether they are based on underlying stereotypes that lead to different ratings. By applying a multi-trait–multimethod model in a latent-variable framework (Eid, 2000; Eid et al., 2003) to parent and teacher ratings of kindergarteners, the present study not only examines the construct validity of three sub-dimensions of social-emotional competence (aggressive behavior, cooperative behavior, and emotional self-regulation) but also whether method effects differ because of sibling status.

Social-Emotional Competence

Social-emotional competence is defined as the *effectiveness in interaction* and develops from early childhood onwards (Denham et al., 2014). It consists of cognitive, emotional, and behavioral aspects (Denham, 2006). Socially and emotionally competent individuals are able to reach their own goals in a socially accepted manner over time and across situations (Kanning, 2002). Social-emotional competence fosters friendships (Barry & Wigfield, 2002), decreases social anxieties, and even reduces the propensity to commit crimes (Greco & Morris, 2005; Stepp et al., 2011). In the educational context, it is positively associated with a person's educational success (Denham et al., 2009, 2014).

Children as young as 3 years face many new tasks that refer to several sub-dimensions of social-emotional competence. They have to integrate themselves into a group of peers, learn new norms and rules of conduct, and make compromises (Denham et al., 2009). Therefore, this study focused on three sub-dimensions of social and emotional competence that develop early and are strongly related to each other: (1) *cooperative behavior* and *(2) (non-)aggressive behavior* are indicators of external social-emotional competence, where one's social behavior is oriented on other's goals and needs. (3) *Emotional self-regulation* refers to internal processes of social-emotional competence such as the awareness, regulation, and expression of own emotions but also the understanding of others' emotions.

Agreement between Informants and Potential Influences of Siblings

Many studies (Dinnebeil et al., 2013; Fält et al., 2018) and metastudies (Achenbach et al., 1987; Renk & Phares, 2004; Rescorla et al., 2012) on social and disruptive behavior of kindergarteners, revealed low correlations between ratings from teachers and parents. These studies showed that

correlations were higher for externalizing than internalizing problem behavior (Rescorla et al., 2012; Winsler & Wallace, 2002). Also, higher informant-correlations has been found for 6–to 11-year-olds than for adolescents or kindergarteners (Achenbach et al., 1987; Renk & Phares, 2004), for girls than for boys (Gagnon et al., 1992; Grills & Ollendick, 2003), and for higher educated mothers as raters (Gagnon et al., 1992). So far, the effect of sibling status on independent ratings has been hardly investigated. To our knowledge, only one study used independent ratings to assess the social-emotional competence of siblings: Downey and Condron (2004) investigated the impact of siblings on social-emotional competence using parent and teacher ratings. They report that kindergarteners with one or two siblings were perceived as more socially competent by the teachers than children with none or more than two siblings. No such differences were found for parent ratings (Downey & Condron, 2004).

Considering the multidimensionality of assessments due to different informants, it is unclear if the sibling status affects children's true social-emotional competence or the informant bias: On the one hand, only children might show higher social-emotional competence than children with siblings or vice versa. On the other hand, informant biases (differences in the method variance) could be moderated because the child for assessment is an only child or a child with siblings. Reasons for this are manifold. First, it is unclear whether parents and teachers were affected by stereotypes regarding their ratings and whether they were influenced equally. Only children were often associated with negative stereotypes such as being selfish, lonely, socially estranged, self-centered, unlikable, or maladjusted (e.g., Polit & Falbo, 1987; Sulloway, 1995; Thompson, 1974). Second, families with more than one child could have more comparison possibilities within their family than families with only one child. The assessments of parents could therefore vary across groups.

Research Questions

Until now, only a few studies have investigated the influence of informant bias (method factor) for assessments of social-emotional competence of kindergarteners (Ferreira et al., 2021; Low et al., 2015; Yu et al., 2015). Particularly, sibling status has been neglected in previous studies as a reason for informant bias. Consequently, we addressed the following questions:

1. Do kindergarten teachers and parents rate children's social-emotional competence comparably, or do informant-specific effects bias the assessment?
2. Are multi-informant assessments of social-emotional competence comparable for only children and children with siblings, or does sibling status affect informant-specific effects?

Method

Participants

The data were part of a larger research project (see Weinert et al., 2013) and originally included 547 three-year-old children from two German federal states (Bavaria, Hesse). Because for 14 children no valid responses were observed, the analyzed sample reduced to $n = 532$ children ($M = 39.2$ months; $SD = 0.2$; 47% girls) that each were evaluated by one kindergarten teacher ($M = 39.4$ years; $SD = 0.5$; 94% women) and one parent ($M = 34.6$ years; $SD = 0.2$; 95% women). About 23% ($n = 120$) of the children had no siblings, while for the rest ($n = 412$) the median was 1 ($min = 1$, $max = 5$) biological, adopted, foster, or stepsibling. 29% ($n = 119$) of children with siblings were firstborns and around 4% ($n = 16$) were multiples (twins). Sociodemographic differences between the two child groups were negligible (see Table S1 in the supplemental material), with a slightly

lower percentage of girls among only children (45% vs. 48%) but a comparable mean age (Cohen's d = 0.09).

Measure

An adaptation of a German short form of the California-Child-Q-Sort (Göttert & Asendorpf, 1989) was used to assess three facets of social-emotional competence, that is, cooperative behavior, aggressive behavior, and emotional self-regulation. Each facet was measured by one parent and one teacher with three items on four-point response scales from 1 = *do not agree at all* to 4 = *totally agree* (see Table 1). Negatively worded items were reverse scored. Because of missing values, there were n = 442 parent ratings and n = 502 teacher ratings available. Descriptive statistics, including means, standard deviations, and correlations between the items of the three scales, are summarized in Table S2 and Table S3 in the supplemental material. McDonald's omega reliabilities in the total sample and the two child group subsamples ranged between ω = .60–.78 for parent ratings and ω =.71–.91 for teacher ratings, thus, indicating acceptable to good reliabilities for all scales (Table 2).

Analytical Approach

The construct validity of the social-emotional competence scales was examined using multi-trait multi-method (MTMM) analyses in a confirmatory factor analytic (CFA) framework. Following Eid and colleagues (2003), we estimated a correlated trait—correlated method minus-one (CTC(M-1)) model that specified three correlated trait factors for cooperative behavior, aggressive behavior, and emotional self-regulation and three correlated method factors for parent ratings (see Figure 1). In this approach, a reference method must be selected based on theoretical assumptions. Teacher ratings were selected as a reference method due to their professional training and daily routine with children. In contrast, the method factors presented unique variances in parent ratings. We modeled different method factors for each subscale to examine the degree of method effects generalized across scales. Trait and method factors were allowed to correlate among themselves but not with each other. From this model, two indices were derived (Eid et al., 2003): (a) the consistency coefficient reflected the part of the variance of the non-reference method indicator explained by the comparison standard, that is, how well differences detected in parent ratings can be predicted by differences in the teacher ratings; (b) the method-specificity coefficient represented the unexplained part of the variance of a non-reference method indicator, that is, the

Table 1. Items of the Social-Emotional Competence Scale.

Facet		Item: "My \ the child …"
Aggressive behavior	A1*	Often is aggressive toward others
	A2*	Often starts arguing and fighting with others
	A3*	Often teases other children
Cooperative behavior	C1	Gets along well with other children
	C2	Is admired and sought out by other children
	C3	Is helpful and cooperative
Emotional self-regulation	E1*	Often reacts exaggeratedly mad when frustrated
	E2*	Gets angry easily
	E3*	My child is easily offended or miffed

Note. * Negatively worded items were reverse scored.

Table 2. McDonald's Omega Reliabilities for Social-Emotional Competence Scales by Informant.

	Total sample		Only children		Children with siblings	
	Parent	Teacher	Parent	Teacher	Parent	Teacher
Aggressive behavior	.72	.90	.68	.91	.73	.89
Cooperative behavior	.60	.73	.64	.71	.60	.73
Emotional self-regulation	.77	.83	.78	.83	.77	.83

Note. n = 532 children in total, n = 120 only children and n = 412 children with siblings.

influence of a specific method, in our case, the parent. Consistency and method-specificity coefficients in a CTC(M-1) model can be calculated for observed and true scores (Eid et al., 2003).

Convergent validity can be inferred if the latent correlations between teacher and parent ratings, that is, the square root of the consistency coefficients (Eid et al., 2003), are large, thus, indicating that parents and teachers rated children similarly. Moreover, consistency coefficients should be larger than the method-specificity coefficient. In contrast, discriminant validity can be inferred if the correlations between the latent trait factors in the CTC(M-1) model are lower than 1.00, that is, $r < .85$ (Brown, 2006). Moreover, correlations between the method factors for different traits show whether the method effects generalize across scales and parents consistently rate differently compared to teachers (Eid et al., 2003).

Differences in the construct validity of the administered scales for only children and children with siblings were studied using multi-group CFAs. We investigated measurement invariance (Steenkamp & Baumgartner, 1998; Vandenberg & Lance, 2000) by comparing increasingly restrictive models. Configural measurement invariance was supported if the MTMM model (without any cross-group constraints) fitted comparably in both child groups (Meredith, 1993), while metric measurement invariance was inferred if cross-group constraints on the factor loadings did not impair the model fit. Scalar measurement invariance with constrained factor loadings and thresholds was also supported.

We used a full information maximum likelihood (Enders, 2010) estimator to handle missing values in all indicator variables (item non-response = 38%). The CFAs were estimated in *Mplus* version 7 (Muthén & Muthén, 1998-2012) with a weighted least square estimator with adjusted mean and variance χ^2 test of model fit (WLSMV estimator; Nussbeck et al., 2006). Model fit was evaluated in line with prevalent standards interpreting comparative fit indices (CFI) $\geq .95$, root mean square errors of approximation (RMSEA) $\leq .08$, and weighted root mean square residuals (WRMR) ≤ 1.0 as "acceptable" and models with CFI $\geq .97$, RMSEA $\leq .05$, and WRMR $\leq .90$ as "good" fitting (DiStefano et al., 2018; Hu & Bentler, 1999). Model comparisons were based on Satorra-Bentler-χ^2-difference tests (Satorra & Bentler, 2001) and differences in CFIs for which values $\leq -.01$ indicated comparable models (Cheung & Rensvold, 2002).

Results

Multi-Trait Multi-Informant Analyses of Social-emotional Competence

The CTC(M-1) model (see Figure 1) for the total sample (n = 532) fitted well to the data, χ^2 (120) = 240.50, $p < .001$; RMSEA = .04, 90% CI [.04, .05]; CFI = .99; WRMR = .89. All teacher ratings had substantial loadings on the latent trait factors, $Mdn(\beta) = 0.62–.98$,
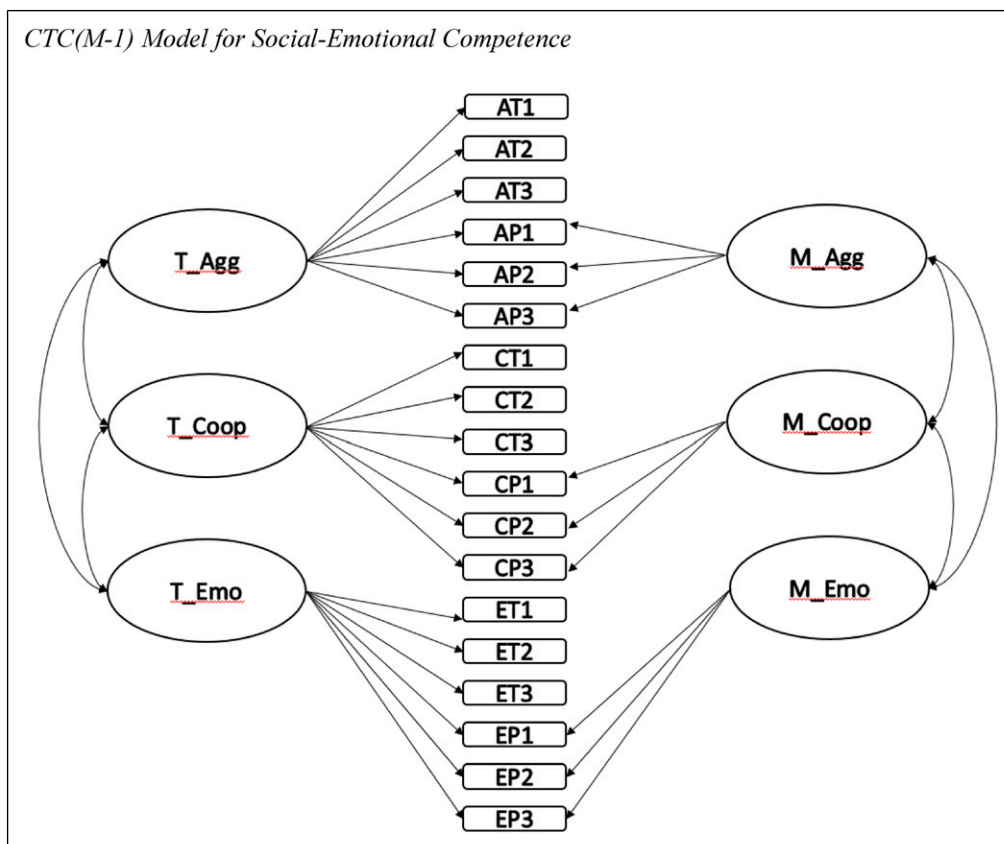
Figure 1. CTC(M-1) Model for Social-Emotional Competence. *Note.* Agg = aggressive behavior, Coop = cooperative behavior, Emo = emotional self-regulation, T = Trait factors, M = Method factors, Teacher and parent rating of aggressive behavior (AT1-3 & AP1-3), cooperative behavior (CT1-3 & CP1-3), and emotional self-regulation (ET1-3 & EP1-3).

whereas the respective parent ratings were markedly lower, $Mdn(\beta) = 0.14–.34$. In contrast, the parent ratings exhibited substantial loadings on the method factors, $Mdn(\beta) = .52–.84$ (Table 3). The different indices derived from this model are summarized in Table 4. Evidence for convergent validities of the three traits was limited. Although the latent correlations between teacher and parent ratings were moderate for most indicators, ranging between .25 and .42, all method-specificity coefficients were substantially larger than the consistency coefficients. Thus, the two informants seemed to measure rather different constructs because differences in teacher ratings cannot accurately predict differences in the parent ratings. In contrast, discriminant validity was generally better supported, as demonstrated by the correlations between the latent trait factors (see Supplemental Table S4). However, the three trait factors correlated significantly ($p < .01$) with each other, $rs = .55$ to .81, the correlations were substantially smaller than our threshold of .85. The method effects also generalized across the different subscales as indicated by the moderate correlations between the method factors, $rs = .29$ to .64. Taking together, these analyses supported discriminant validity between traits but indicated only limited convergent validity across informants.

Table 3. Standardized Loading Parameters of the CTC(M-1)Model for the Total Sample and the Child Group Subsamples.

Item	Total sample		Only children		Children with siblings	
	Trait loading	Method loading	Trait loading	Method loading	Trait loading	Method loading
Aggressive behavior						
AT1	.91		.91		.91	
AT2	.93		.93		.93	
AT3	.91		.90		.91	
AP1	.29	.63	.54	.57	.23	.66
AP2	.34	.74	.42	.71	.32	.74
AP3	.21	.72	.18	.66	.21	.74
Cooperative behavior						
CT1	.98		.96		.96	
CT2	.62		.59		.64	
CT3	.66		.63		.66	
CP1	.29	.74	.51	.67	.29	.74
CP2	.20	.61	.39	.50	.20	.63
CP3	.14	.52	.37	.45	.11	.49
Emotional self-regulation						
ET1	.78		.77		.77	
ET2	.99		.98		1.00	
ET3	.74		.70		.75	
EP1	.28	.68	.18	.70	.30	.68
EP2	.29	.84	.19	.98	.32	.79
EP3	.20	.67	.18	.58	.21	.69

Note. Teacher and parent rating of aggressive behavior (AT1-3 & AP1-3), cooperative behavior (CT1-3 & CP1-3), and emotional self-regulation (ET1-3 & EP1-3). Blank cells indicate factor loadings fixed to zero by definition of the model. CT-C(M−1) = correlated trait–correlated method minus one model.

The Role of Sibling Status

The model fit for different levels of measurement invariance between only children and children with siblings is summarized in Table 5. The unconstrained model (Table 5, M1) resulted in a satisfactory fit, χ^2 (239) = 376.02, p < .001; RMSEA = .05, 90% CI [.04, .06]; CFI = .98; WRMR = 1.094, indicating comparable factor structures in both groups of children. Constraining the loadings on the factor loadings (Model M2) and additionally, the thresholds (Model M3) did not result in a loss of fit as indicated by non-significant difference tests (p > .05) and differences in CFIs <.01. Finally, we examined convergent and discriminant validity between child groups. Group results are comparable to total sample results indicating low convergent validity and discriminant validity between traits (Table 6 & Supplemental Table S4). Latent correlations between parent and teacher assessments of emotional self-regulation (.29–.40) were higher for children with siblings. However, latent correlations for assessments of cooperative behavior (.60–.63) and aggressive behavior (.27–.69) were higher for only children (Table 4). Similar large method-specificity coefficients (.60–.96; except AP1 for only children: .53) in both child groups indicate low convergent validity. Group comparison showed further that 7–47% of parent ratings on aggressive behavior, 37–40% of parent ratings on cooperative behavior, and 4–9% of parent

Table 4. Convergent and Discriminant Validity.

	Total sample			Only children			Children with siblings		
	Corr	Con	Spec	Corr	Con	Spec	Corr	Con	Spec
AT1	1.00	1.00		1.00	1.00		1.00	1.00	
AT2	1.00	1.00		1.00	1.00		1.00	1.00	
AT3	1.00	1.00		1.00	1.00		1.00	1.00	
AP1	.42	.18	.82	.69	.47	.53	.32	.10	.90
AP2	.42	.18	.82	.51	.26	.74	.39	.15	.85
AP3	.28	.08	.92	.27	.07	.93	.28	.08	.92
CT1	1.00	1.00		1.00	1.00		1.00	1.00	
CT2	1.00	1.00		1.00	.89		1.00	1.00	
CT3	1.00	1.00		1.00	.91		1.00	1.00	
CP1	.36	.13	.87	.60	.37	.63	.36	.13	.87
CP2	.30	.09	.91	.61	.37	.63	.30	.09	.91
CP3	.25	.06	.94	.63	.40	.60	.22	.05	.95
ET1	1.00	1.00		1.00	1.00		1.00	1.00	
ET2	1.00	1.00		1.00	1.00		1.00	1.00	
ET3	1.00	1.00		1.00	1.00		1.00	1.00	
EP1	.38	.15	.85	.25	.06	.94	.40	.16	.84
EP2	.33	.11	.89	.19	.04	.96	.37	.14	.86
EP3	.28	.08	.92	.29	.09	.91	.29	.08	.92

Note. Teacher and parent rating of aggressive behavior (AT1-3 & AP1-3), cooperative behavior (CT1-3 & CP1-3), and emotional self-regulation (ET1-3 & EP1-3). n = 532. Con. = Consistency, Spec. = Method specificity, Corr. = Latent correlation with the standard method ($\sqrt{\text{consistency}}$).

Table 5. Model Fit of the Multi-Group CTC(M-1) Model.

Model	χ^2 (df)	S-B $\Delta\chi^2$ (Δdf)	RMSEA 90% CI	WRMR	CFI	ΔCFI (ref. M1)
M1 Unconstrained baseline model	376.02* (239)	-	.05 [.04, .06]	1.09	.98	-
M2 M1 + factor loadings constrained	383.33* (254)	23.85 (15)	.04 [.04, .05]	1.20	.98	< −.01
M3 M1 + factor loadings and thresholds constrained	413.46* (296)	30.8 (42)	.04 [.03, .05]	1.23	.98	< −.01

Note. n = 120 only children and n = 412 children with siblings. * significant at $p < .01$.

ratings on emotional self-regulation of only children could be explained by teacher ratings. In contrast, only 8–15%, 5–13%, and 8–16%, respectively, of parent ratings for children with siblings could be explained by teacher ratings.

Given threshold invariance was supported (see Table 5), we also examined mean level differences. Only children were evaluated significantly lower on aggressive behavior, $d = .35$, $p < .05$, emotional self-regulation, $d = .51$, $p < .05$, and co-operative behavior $d = .32$, $p < .05$. Moreover, all three method factors showed no significant ($p > .05$) mean-level differences.

Table 6. Correlations of the Trait and Method Factors in the CTC(M − 1) Model for Child Group Subsamples.

	Trait factors			Method factors		
	Aggr. behavior	Coop. behavior	Emotional self-regulation	Aggr. behavior	Coop. behavior	Emotional self-regulation
Trait Factors						
Agg. b.	-	.57*	.85*			
Coop. b.	.67*	-	.41*			
Emo. s.-reg	.80*	.61*	-			
Method Factors						
Agg. b.				-	.69*	.67*
Coop. b.				.69*	-	.51*
Emo. s.-reg				.62*	.30*	-

Note. Correlations from the sibling's sample are shown below the diagonal; correlations from the only child sample are shown above the diagonal; n = 120 only children and n = 412 children with siblings. Variance fixed to 1. * significant at $p < .01$.

Discussion

Previous research has shown that the correlation between children's social-emotional competence ratings collected from two different informants is often relatively low (Dinnebeil et al., 2013; Rescorla et al., 2012). Reasons for these results could be different interpretations between informants, differences in social settings, or different thresholds for identifying behavior (De los Reyes et al., 2013). Still, only a few studies investigated these low correlations further (Ferreira et al., 2021; Low et al., 2015). In our MTMM analyses, it was possible to examine multi-informant assessments in more detail, which has not been done often to investigate assessments of children's social-emotional competence (Ferreira et al., 2021; Gomez, 2014; Low et al., 2015; Yu et al., 2015). The present study led to three central findings. First, the administered scales exhibited rather good reliabilities, and metric invariance between child groups was supported. Second, the MTMM model demonstrated discriminant validity between traits in line with other studies (Ferreira et al., 2021; Gomez, 2014) but only limited convergent validity across informants. Third, despite a whole range of stereotypes about only children, sibling status has, so far, not been considered a reason for low agreement between informants. However, independent analyses for only children and children with siblings lead to similar results as for the total sample with adequate discriminant validity between traits but low convergent validity.

Our findings suggest that parents and teachers do indeed rate children's competence differently. This could be due to different social contexts. Further analyses with different rating scales and more ratings from both social contexts would be needed to test these assumptions thoroughly. Until then, an important decision on the child's future should not rely on one perspective only. Furthermore, results show no significant difference between ratings of only children and children with siblings. Only descriptive results indicate that parents with one or more than one child might apply different thresholds for the child's competence. They, therefore, agree more or less with the teacher rating on each sub-dimension. For example, parent ratings of cooperative behavior of only children are more similar to teacher ratings. One reason for that might be that parents and teachers refer to the same or a similar social context where the child encounters peers. Another reason might be that scale

properties such as positively versus negatively worded items lead to higher consistency between teachers and parents of only children. Finally, about one third of children with siblings were firstborns. These children are likely to resemble only children in that they were temporary only children, especially at the age of three when a sibling has just arrived.

Limitations

This research was not without limitations. The relatively small sample prohibited more detailed analyses based on this data, such as the effects of the number of siblings, birth order, or birth spacing on the ratings of social-emotional competence of parents and teachers. One shortcoming of the method is that results are not symmetrical. By changing the reference method from teacher rating to parent rating, fit indices might be different (Eid et al., 2003). However, secondary analyses with a changed reference method confirmed the robustness of our main results (see Tables S5–S8 in the supplemental material).

Conclusion

Our results indicate that kindergarten teachers and parents do not rate the competence of children comparably. This could be attributed to the fact that the informants observe children in different social situations. Additionally, our results show no general differences in regard to sibling status. However, descriptive results indicate minor differences in the comparability of parent and teacher ratings by sibling status as a function of the sub-dimensions of social-emotional competence. Therefore, sibling status explains only a small, not significant, part of the low agreement between informants.

Declaration of Conflicting Interests

Funding

Data Availability

The data and material are available at https://doi.org/10.5159/IQB_BIKS_3_10_v6.

ORCID iDs

Carina Schönmoser https://orcid.org/0000-0002-8825-1095
Claudia Karwath https://orcid.org/0000-0002-2278-6409
Timo Gnambs https://orcid.org/0000-0002-6984-1276

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioural and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213–232. https://doi.org/10.1037/0033-2909.101.2.213.

Barry, C. M., & Wigfield, A. (2002). Self-perceptions of friendship-making ability and perceptions of friends' deviant behavior: Childhood to adolescence. *The Journal of Early Adolescence*, *22*(2), 143–172. https://doi.org/10.1177/0272431602022002002.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford Press. https://doi.org/10.1037/h0046016.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5.

De Los Reyes, A., Thomas, S., Goodman, K., & Kundey, S. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, 9(1), 123–149. https://doi.org/10.1146/annurev-clinpsy-050212-185617.

Denham, S. A. (2006). Social-emotional competence as support for school readiness: what is it and how do we assess it? *Early Education and Development*, 17(1), 57–89. https://doi.org/10.1207/s15566935eed1701_4.

Denham, S. A., Bassett, H. H., Zinsser, K., & Wyatt, T. M. (2014). How preschoolers' social-emotional learning predicts their early school success. Developing theory-promoting, competency-based assessments. *Infant and Child Development*, 23(4), 426–454. https://doi.org/10.1002/icd.1840.

Denham, S. A., Wyatt, T. M., Bassett, H. H., Echeverria, D., & Knox, S. S. (2009). Assessing social-emotional development in children from a longitudinal perspective. *Journal of Epidemiological Community Health*, 63(Suppl 1), i37–52. https://doi.org/10.1136/jech.2007.070797.

Dinnebeil, L. A., Sawyer, B. E., Logan, J., Dynia, J. M., Cancio, E., & Justice, L. M. (2013). Influences on the congruence between parents' and teachers' ratings of young children's social skills and problem behaviors. *Early Childhood Research Quarterly*, 28(1), 144–152. https://doi.org/10.1016/j.ecresq.2012.03.001.

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal*, 25(3), 453–466. https://10.1080/10705511.2017.1390394.

Downey, D. B., & Condron, D. J. (2004). Playing well with others in kindergarten. The benefit of siblings at home. *Journal of Marriage and the Family*, 66(2), 333–350. https://doi.org/10.1111/j.1741-3737.2004.00024.x.

Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, 65(2), 241–261. https://doi.org/10.1007/BF02294377.

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C (M-1) model. *Psychological Methods*, 8(1), 38–60. https://doi.org/10.1037/1082-989X.8.1.38.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Fält, E., Wallby, T., Sarkadi, A., Salari, R., & Fabian, H. (2018). Agreement between mothers', fathers', and teachers' ratings of behavioural and emotional problems in 3-5-year-old children. *Plos One*, 13(11), Article e0206752. https://doi.org/10.1371/journal.pone.0206752.

Ferreira, T., Geiser, C., Cadima, J., Matias, M., Leal, T., & Mena Matos, P. (2021). The strengths and difficulties questionnaire: An examination of factorial, convergent, and discriminant validity using multitrait-multirater data. *Psychological Assessment*, 33(1), 45–59. https://doi.org/10.1037/pas0000961.

Gagnon, C., Vitaro, F., & Tremblay, R. E. (1992). Parent-teacher agreement on kindergarteners' behavior problems: A research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 33(7), 1255–1261. https://doi.org/10.1111/j.1469-7610.1992.tb00944.x.

Gomez, R. (2014). Correlated trait–correlated method minus one analysis of the convergent and discriminant validities of the Strengths and Difficulties Questionnaire. *Assessment*, 21(3), 372–382. https://doi.org/10.1177/1073191112457588.

Göttert, R., & Asendorpf, J. (1989). Eine deutsche version des California-child-Q-sort (Block & Block, 1980). Kurzform. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 21(1), 70–82.

Greco, L. A., & Morris, T. L. (2005). Factors influencing the link between social anxiety and peer acceptance: Contributions of social skills and close friendships during middle childhood. *Behavior Therapy*, *36*(2), 197–205. https://doi.org/10.1016/S0005-7894(05)80068-1.

Grills, A. E., & Ollendick, T. H. (2003). Multiple informant agreement and the anxiety disorders interview schedule for parents and children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(1), 30–40. https://doi.org/10.1097/00004583-200301000-00008.

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

Kanning, U. P. (2002). Soziale Kompetenz – Definition, Strukturen und Prozesse. *Zeitschrift für Psychologie*, *210*(4), 154–163. https://doi.org/10.1026//0044-3409.210.4.154.

Low, J. A., Keith, T. Z., & Jensen, M. (2015). What predicts method effects in child behavior ratings. *Journal of Psychoeducational Assessment*, *33*(2), 177–187. http://dx.doi.org/10.1177/0734282914544922.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. https://doi.org/10.1007/BF02294825.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide* (7th ed.). Muthén & Muthén.

Nussbeck, F. W., Eid, M., & Lischetzke, T. (2006). Analysing multitrait-multimethod data with structural equation models for ordinal variables applying the WLSMV estimator: Wat sample size is needed for valid results? *The British Journal of Mathematical and Statistical Psychology*, *59*(1), 195–213. https://doi.org/10.1348/000711005X67490.

Polit, D. F. & Falbo, T. (1987). Only children and personality development: A quantitative review. *Journal of Marriage and the Family*, *49*(2), 309–325. https://doi.org/10.2307/352302.

Renk, K., & Phares, V. (2004). Cross-informant ratings of social competence in children and adolescents. *Clinical Psychology Review*, *24*(2), 239–254. http://dx.doi.org/10.1016/j.cpr.2004.01.004.

Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Bilenberg, N., Bjarnadottir, G., Denner, S., Dias, P., Dobrean, A., Döpfner, M., Frigerio, A., Gonçalves, M., Guðmundsson, H., Jusiene, R., Kristensen, S., Lecannelier, F., Leung, P. W. L., Liu, J., Löbel, S. P., Machado, B. C., & Verhulst, F. C. (2012). Behavioral/emotional problems of preschoolers: Caregiver/teacher reports from 15 societies. *Journal of Emotional and Behavioral Disorders*, *20*(2), 68–81. https://doi.org/10.1177/1063426611434158.

Satorra, A. & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507–514. https://doi.org/10.1007/BF02296192.

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*(1), 78–107. https://doi.org/10.1086/209528.

Stepp, S. D., Pardini, D. A., Loeber, R., & Morris, N. A. (2011). The relation between adolescent social competence and young adult delinquency and educational attainment among at-risk youth: The mediating role of peer delinquency. *Canadian Journal of Psychiatry*, *56*(8), 457–465. https://doi.org/10.1177/070674371105600803.

Sulloway, F. (1995). Birth order and evolutionary psychology: A meta-analytic overview. *Psychological Inquiry*, *6*(1), 75–80. https://doi.org/10.1207/s15327965pli0601_15.

Thompson, V. D. (1974). Family size: Implicit policies and assumed psychological outcomes. *Journal of Social Issues*, *30*(4), 93–124. https://doi.org/10.1111/j.1540-4560.1974.tb01757.x.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–69. https://doi.org/10.1177/109442810031002.

Weinert, S., Roßbach, H.-G., Faust, G., Blossfeld, H.-P., & Artelt, C. (2013). *Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter. (BiKS-3-10)* (Version 6) [Datensatz]. IQB. http://doi.org/10.5159/IQB_BIKS_3_10_v6.

Winsler, A., & Wallace, G. (2002). Behavior problems and social skills in preschool children: Parent-teacher agreement and relations with classroom observations. *Early Education and Development*, *13*(1), 41–58. http://doi.org/10.1207/s15566935eed1301_3.

Yu, J., Sun, S., & Cheah, C. S. (2015). Multitrait-multimethod analysis of the strengths and difficulties questionnaire in young Asian American children. *Assessment*, *23*(5), 603–613. https://doi.org/10.1177/1073191115586459.