Pawel R. Kulawiak | Karolina Urton | Johanna Krull | Thomas Hennemann | Jürgen Wilbert

Internalizing Behavior of Sociometrically Neglected Students in Inclusive Primary Classrooms

A Methodological Issue?

Check for
updates

Internalizing Behavior of Sociometrically Neglected Students in Inclusive Primary Classrooms – A Methodological Issue?

Pawel R. Kulawiak[1,2]*, Karolina Urton[2], Johanna Krull[2], Thomas Hennemann[2] and Jürgen Wilbert[1]

[1] Department of Inclusive Education, University of Potsdam, Potsdam, Germany, [2] Department of Special Education and Rehabilitation, University of Cologne, Cologne, Germany

Internalizing problems in children belong to the category of special educational needs called emotional and behavioral difficulties. Recent decades have witnessed a critical discussion about whether children and adolescents experiencing internalizing problems are at risk of being sociometrically neglected (neither liked nor disliked by their peers). Previous studies have shown evidence both for and against the association between internalizing problems and neglected sociometric status. These contradictory results may be due to the following methodological aspects: (1) shortcomings of sociometric status classification methods (arbitrariness of the sociometric classification rules) and (2) different operationalizations of internalizing problems (broadband and narrowband dimensions of behavior). The aim of the present study is to investigate empirically whether and to what extent these methodological aspects lead to contradictory results on the internalizing behavior of neglected students. This question is investigated using a sample of students ($N = 2334$) in German inclusive primary schools. The systematic investigation presented here provides initial indications that the various methodological approaches can lead to conflicting results. The contradictory results are not only due to the application of different sociometric classification methods, but also to different operationalizations of internalizing behavior (narrowband and broadband scales). Earlier contradictory evidence on the internalizing behavior of neglected students must therefore be seen in a different light: the reasons for previously conflicting results may actually be methodological. Based on the results, conclusions are drawn as to how methodological aspects can be given more consideration in sociometric research on internalizing behavior.

Keywords: special educational needs, inclusive education, social inclusion, sociometric status, sociometric neglect, internalizing behavior, broadband and narrowband dimensions of behavior

INTRODUCTION: SOCIAL INCLUSION OF STUDENTS WITH SPECIAL EDUCATIONAL NEEDS (SEN)

In view of the educational placement of students with SEN in mainstream schools, a major challenge and goal for education policy and educators is the successful social inclusion of these children. The importance of the topic has been emphasized in previous studies, where it is assumed that students with SEN are an at-risk group for social exclusion in mainstream schools (Frostad et al., 2011;

Krull et al., 2014; Garrote et al., 2017; Henke et al., 2017). However, proper conceptualization and measurement of social inclusion of students with SEN is the subject of an ongoing debate (Chambers and Kay, 1992; Frederickson et al., 2007; Koster et al., 2009; Gerullis and Huber, 2018).

Poor Conceptualization and Measurement of Social Inclusion

Despite the current evidence, previous reviews (Chambers and Kay, 1992; Nakken and Pijl, 2002; Frederickson et al., 2007) have concluded that research results give a contradictory picture of whether students with and without SEN are socially included to equal extents in mainstream classrooms. A reason for the conflicting evidence could be the unidimensional assessment of social inclusion (Chambers and Kay, 1992). Depending on which dimension of social inclusion is addressed and which measurement technique is used, the answer to the question of whether students with SEN are equally socially integrated can vary. Accordingly, the study by Avramidis et al. (2018) considered several dimensions of social inclusion. A differentiated picture emerges here insofar as students with SEN (compared to students without SEN) received fewer nominations of peer acceptance and had fewer friends and social interactions with classmates (sociometric measures), whereas they showed no difference in social self-perception and perception of friendship quality (self-report questionnaires). Furthermore, a comparison of the measurement of social inclusion with two different sociometric tools, i.e., the peer nomination method and social cognitive mapping (Avramidis et al., 2017), indicated that research using sociometric nominations has mainly revealed an unfavorable social inclusion of children with SEN while social cognitive mapping has yielded mixed and positive results. The same is true if sociometric tools are contrasted with observational methods. The use of sociometric tools implies mostly negative effects of social inclusion for students with SEN and the application of observational methods indicates mainly positive effects (Chambers and Kay, 1992). These contradictory findings may be due to the different operationalizations of social inclusion. Koster et al. (2009), who point out the multidimensional nature of social inclusion of students with SEN (e.g., interactions, friendships, self-perception of peer-acceptance, etc.), describe the poor conceptualization and measurement (see also Gerullis and Huber, 2018) to the effect that previous studies deal with the concept of social inclusion (without further specifying it), but use different operationalizations and therefore refer to different dimensions of social inclusion. Hence, inconsistent results could be an effect of different operationalizations. Another reason for the contradictions, which is examined in more detail below, could be that students with SEN are regarded as a homogeneous group.

Students With SEN Are Not a Homogeneous Group

Another reason for the conflicting research findings could be that students with SEN are considered as a homogenous group. Many studies on social inclusion of students with SEN

(e.g., Frostad et al., 2011) do not distinguish between different types of disabilities or difficulties (e.g., mental disability, physical disability, behavior difficulties, and learning difficulties). The educational needs of students with different types of disabilities or difficulties vary considerably, and therefore inconsistent research findings may reflect the fact that students with SEN are not a homogeneous group (O'Mara et al., 2012). Thus, the study by Krull et al. (2014) showed that the social inclusion of students with SEN depends on the type of disability or difficulty. The results indicate that students with both behavior problems and learning difficulties are more likely to be rejected by their peers than students without SEN, although the rejection rate for students with behavior problems is twice as high as for those with learning difficulties. The highest risk of rejection is for students with combined behavior and learning difficulties. Accordingly, it becomes clear that the evidence found on the social inclusion of students with specific SEN cannot simply be transferred to students with other types of disabilities or difficulties. This issue is described in more detail below with reference to students with emotional and behavioral difficulties.

Students With Emotional and Behavioral Difficulties

Even when research on social inclusion focuses on a particular group of students with SEN, namely students with emotional and behavioral difficulties, generalizations may be flawed. As the category of emotional and behavioral difficulties includes students with both internalizing and externalizing problems, a homogeneous sample cannot be assumed even within this group (Kershaw and Sonuga-Barke, 1998). Accordingly, the educational needs of the students vary depending on the behavior at hand (Kershaw and Sonuga-Barke, 1998). This difference also appears as a function of the type of behavior (internalizing vs. externalizing) for social inclusion (Rytioja et al., 2019). The results of the study by Rytioja et al. (2019) imply that externalizing problems are most strongly associated with the controversial sociometric status (students both liked and disliked by many peers) followed by the rejected sociometric status (students liked by few and disliked by many peers), while for students with internalizing problems, the opposite is the case. Thus, internalizing problems are most strongly related to the rejected sociometric status, followed by the controversial status. As far as the popular sociometric status (students liked by many and disliked by few peers) is concerned, externalizing behavior for this group tends to be higher than internalizing behavior, whereas externalizing and internalizing problems are equally pronounced for sociometrically neglected students (neither liked nor disliked by their peers). Overall, the study by Rytioja et al. (2019) shows that internalizing and externalizing problems are related to different qualities of social inclusion (different sociometric groups). However, as research results on the social inclusion of children with internalizing problems are far from clear, especially when the sociometric neglect is used as an indicator of social inclusion, the present study focuses on the group of students with internalizing problems.

Internalizing Behavior and Sociometric Neglect

A positive association between internalizing behavior and sociometric neglect is often assumed (Wilmshurst, 2017, p. 307): *"Neglected children were shy and withdrawn (Coie and Kupersmidt, 1983), had higher levels of social anxiety and were at greater risk for internalizing problems then their non-neglected peers."* However, a closer examination of the evidence shows that the association is not as clear as suggested. While Coie and Kupersmidt (1983) found that neglected children exhibit more shyness, Cantrell and Prinz (1985) report that there is no association between sociometric neglect and shyness, withdrawal, and anxiety. On the contrary, other study results support the association of sociometric neglect with anxiety (La Greca et al., 1988; Strauss et al., 1988). Regarding depression, neglected children are more likely to report depressive feelings (Kupersmidt and Patterson, 1991) and show severe anhedonic symptoms (Hecht et al., 1998). Again, contrary findings can be identified. Burton and Krantz (1990) conclude that neglected children are not prone to depression or anxiety. Meta-analytical results (Newcomb et al., 1993) show that neglected children evidenced more withdrawal and less depression than average children (but rejected children are most at risk). Mega-analytic results (Crews et al., 2007) show that the controversial status is the strongest sociometric risk factor for internalizing behavior, followed by the neglected and rejected status. More recent studies show an increased level of anxiety, depression, loneliness, and withdrawal in neglected children (DeRosier and Thomas, 2003; Friedman, 2004; Kaya, 2007; Woodhouse et al., 2012) while other studies show no increased association between sociometric neglect and anxiety or withdrawal (Delgado et al., 2016; García Bacete and Cillessen, 2017). The latest study (Rytioja et al., 2019) shows that rejected children exhibit the strongest internalizing behavior.

Consequently, there is evidence for and against the association of internalizing behavior with sociometric neglect. Therefore, in recent decades a critical discussion has taken place about whether children experiencing internalizing problems are at risk for being neglected (Hymel and Rubin, 1985; Bierman, 1987; Rubin et al., 1989; Howe, 2010; Kingery et al., 2010; Epkins and Heckler, 2011; Kulawiak and Wilbert, 2019). The findings of about half of the studies presented suggest that internalizing problems are not related to sociometric neglect or that other sociometric groups are more strongly associated with internalizing problems, which seems to be the case especially for the rejected and controversial groups. The evidence on the internalizing behavior of neglected children is therefore best described as contradictory and must be treated with caution (Howe, 2010).

METHODOLOGICAL REASONS FOR THE CONFLICTING FINDINGS ON INTERNALIZING BEHAVIOR IN NEGLECTED CHILDREN

The aforementioned conflicting findings raise the question of the reasons for this phenomenon, which make it possible to better understand the differences in the internalizing behavior of neglected children. Two methodological aspects are discussed in more detail in the following: (1) sociometric status classification issues and (2) different operationalizations of internalizing behavior.

Sociometric Status Classification Issues

A variety of different sociometric status classification methods is still widely used in research on the social inclusion of children with disabilities and behavioral difficulties (Krull et al., 2014; Avramidis et al., 2017; Garrote, 2017; Rytioja et al., 2019). According to Rubin et al. (1989), the diversity of applied sociometric classification rules can be seen as a reason for the equivocal evidence on the behavior of neglected children. Sociometric classification criteria are arbitrarily defined across sociometric classification methods (Terry and Coie, 1991; Mayeux et al., 2007; Kulawiak and Wilbert, 2019), which is particularly apparent for the group of neglected students (Rubin et al., 1989, p. 96):

> *"[...] the standardized score approach popularized by Coie and Dodge and their colleagues [...] has been most often employed in recent sociometric research, yet the specific criteria used to identify sociometric subgroups differs from one research report to the next. This is particularly true in the case of neglected children where criteria appear to vary, in part, as a result of efforts to increase sample size. [...] Minor variations of this sort produce very different frequencies of identified neglected children."*

The arbitrary variation of classification rules (as described by Rubin et al., 1989) is a common but doubtful practice and a methodological problem (probably linked to equivocal research results). As a result, there is low classification agreement across methods (McMullen et al., 2014). Different methods then classify a child into different groups. In view of the arbitrary classification rules, it can be stated that there is no consensus method (Rubin et al., 1989) and no consensus about the state of sociometric neglect (Kulawiak and Wilbert, 2019). McMullen et al.'s (2014) findings support Rubin et al.'s (1989) assumption that the diversity of sociometric classification rules could be a reason for contradictory research results. They compared eight different sociometric classification methods by analyzing the associations between the different sociometric statuses and social withdrawal (assessed via peer nomination). In four methods, the neglected sociometric status correlated most strongly with social withdrawal. In the other four methods, the rejected sociometric status correlated most strongly with social withdrawal. Hence, different sociometric classification methods lead to different results regarding the association of sociometric status with social withdrawal. These results suggest that the equivocal evidence on the sociometric status of students with internalizing problems should also be discussed against the background of the sociometric classification methods used.
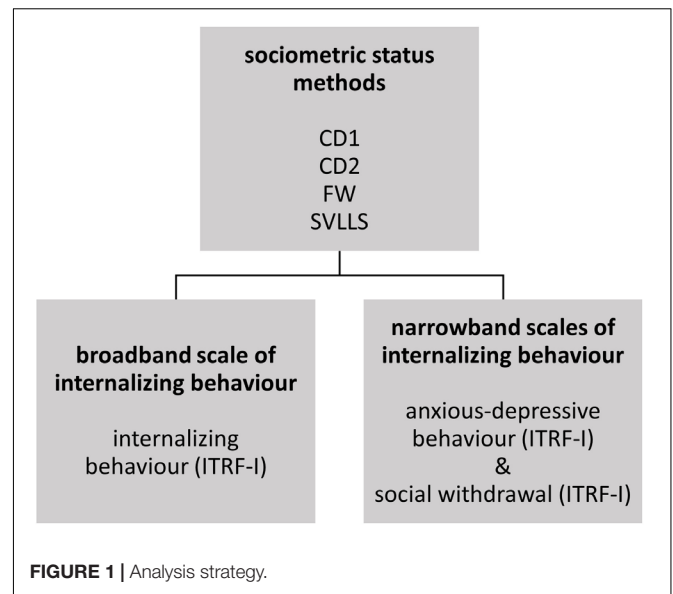
Different Operationalizations of Internalizing Behavior

The second point that should be considered as a reason for the conflicting evidence on internalizing behavior in

neglected children has to do with different operationalizations of internalizing behavior. The behavior of students with internalizing problems is characterized by depressive, anxious, and somatic symptoms, as well as social withdrawal, shyness, sadness, and low self-esteem (Whitcomb, 2018). In line with the current discussion on whether children's behavior can best be described by broadband (internalizing and externalizing behavior) or narrowband categories (e.g., depression, anxiety, and somatization as underlying dimensions of internalizing behavior) (Tandon et al., 2009; Achenbach et al., 2016), it seems equally important to discuss the results of sociometric research against this background. In many cases, broadband scales of behavior are a subsumption of narrowband scales of behavior (Goodman et al., 2010). In earlier sociometric studies, both broadband (e.g., Rytioja et al., 2019) and narrowband scales of behavior (e.g., Tani and Schneider, 1997) were used to asses internalizing behavior. Studies that examine narrowband dimensions of internalizing problems provide an indication that only certain dimensions of internalizing behavior are related to sociometric neglect (differentiated effects). For example, the study by Tani and Schneider (1997) shows that neglected children have a higher degree of somatization than rejected and average children, while depressive symptoms are weaker and feelings of anxiety comparable. Such differentiated effects are also reported in a study by Kulawiak et al. (under review) on the association of narrowband and broadband dimensions of behavior with school-relevant outcomes. Hence, the authors conclude that narrowband scales are more informative than broadband scales of behavior when describing the association between behavior and students' social and academic outcomes. This is also reflected in the fact that some narrowband scales show stronger effect sizes than broadband scales. Accordingly, the question arises as to whether the examination of internalizing behavior in neglected children by a broadband scale does not systematically underestimate the association, whereas the application of narrowband scales would show differentiated associations and stronger effect sizes. Hence, the equivocal evidence on the internalizing behavior of neglected students should also be discussed against the background of the measured dimensions of internalizing behavior (broadband and narrowband). However, it must also be taken into account that a wide range of different measuring approaches is used to assess internalizing behavior (peer ratings, teacher ratings, clinical assessments, etc.). This variety can also impact research results on internalizing behavior in neglected children.

AIMS

The considerations presented about the methodological factors, which can determine the research results on the internalizing behavior of sociometrically neglected students, emphasize the importance of a systematic investigation of these methodological aspects in order to gain a clear understanding of previously conflicting findings. Accordingly, the present data on internalizing behavior in neglected students are considered from a methodological perspective (**Figure 1**): (1) the application of different sociometric status classification methods and (2) the

FIGURE 1 | Analysis strategy.

measurement of internalizing problems by means of broadband and narrowband scales of behavior.

With regard to the first methodological aspect, it is evaluated whether different sociometric classification methods give contradictory results in the prediction of internalizing behavior. The rank order of sociometric groups in terms of level of internalizing behavior is crucial for judging whether the results of the different sociometric classification methods are contradictory. If all sociometric methods show the same rank order of sociometric groups (e.g., popular, average, controversial, rejected, and neglected), the results are consistent. If the rank orders differ (e.g., popular, average, controversial, rejected, and neglected vs. popular, average, controversial, neglected, and rejected), the results are contradictory. If the rank of the neglected group differs across rank orders (i.e., across methods), the results are contradictory with regard to the neglected children.

With regard to the second methodological aspect, internalizing problems are assessed using both narrowband and broadband scales of behavior. It is assumed that the use of narrowband scales shows differentiated effect sizes and that the use of the broadband scale underestimates the association between sociometric neglect and internalizing behavior (i.e., the narrowband scales show stronger effect sizes).

Accordingly, the aim of the present study is therefore to empirically investigate whether and to what extent the methodological aspects contribute to the generation of contradictory research results and, if so, to draw conclusions about how to overcome these issues in future research on the internalizing behavior of neglected children.

MATERIALS AND METHODS

Procedure and Participants
The present study is part of the German research project "Schools On Their Journey Towards Inclusion: Mettmann 2.0"

(Hennemann et al., 2018). Data were collected in 2017 and 2018 at ten inclusive primary schools in Mettmann County, Federal State of North Rhine-Westphalia, Germany. As the study was approved by the local school authority of Mettmann County (approval criteria: compliance with data protection regulations and educational relevance of research), additional ethics approval was not required in accordance with the national legislation and the institutional requirements. Written and informed consent was obtained from the children's parents/legal guardians. The sample consists of 112 classes across grades one to four, i.e., 12 (11%) classes of grade one, 19 (17%) classes of grade two, 25 (22%) classes of grade three, 27 (24%) classes of grade four, and 29 (26%) multigrade classes of grades one to four (different grade compositions). The median class size is 24 ($Min = 16$, $Max = 32$). The sample comprises 2,699 students in total, but complete data is available for 2,334 students ($M_{age} = 8.81$ years, $SD_{age} = 1.20$ years, $N_{boys} = 1205$ [52%]). Missing data are due to lack of parental consent or the absence of students at data collection, e.g., due to illness. For more than half of the classes, complete data are available on more than 90% of the children in a class (proportion of complete data per class: $Min = 43\%$, $Mdn = 91\%$, $Max = 100\%$).

Measures

Sociometry

To collect sociometric data, a sociometric nomination questionnaire was used (Bukowski et al., 2012). Second to fourth-grade students filled out the questionnaire on their own in the classroom. Children in the first grade who could not sufficiently write and read as well as children who needed special support were questioned in a one-to-one interview in a separate room. All children were instructed to write down the names of their classmates (no self-nomination) whom they *liked the most* (social acceptance) and whom they *liked the least* (social rejection). The number of nominations was unlimited. Nominations received for each question (indegrees) are counted for each student, resulting in two scores (LM: like-most score, LL: like-least score). An acceptable reliability of sociometric data ($\alpha > 0.80$) can be shown even with low participation rates (participation rate per class > 10%) (Marks et al., 2013). Four different methods were used to determine students' sociometric status (sociometric groups: average, popular,[1] rejected, neglected, and controversial).

Method 1 (CD1)

Using the classification procedure by Coie and Dodge (1983) (hereinafter referred to as CD1), students are classified into the different sociometric status groups on the basis of within-class standardized scores (LM_z and LL_z; $M = 0$ and $SD = 1$ for each class). Additionally, a social preference score (SP_z: within-class standardized difference between LM_z and LL_z) and a social impact score (SI_z: within-class standardized sum of LM_z and LL_z) are required for sociometric group definition.

Classification rules are subsequently applied to assign children to the groups:

- Popular: $LM_z > 0$ and $LL_z < 0$ and $SP_z > +1$
- Rejected: $LM_z < 0$ and $LL_z > 0$ and $SP_z < -1$
- Neglected: $LM_z < 0$ and $LL_z < 0$ and $SI_z < -1$
- Controversial: $LM_z > 0$ and $LL_z > 0$ and $SI_z > +1$

Coie and Dodge (1983) advocate including all remaining students (who could not be assigned to a particular group by the abovementioned classification rules) in the average group.

Method 2 (CD2)

Other authors (Boivin et al., 1994; Hubbard, 2001) have adjusted the classification rules of the CD1 method and use ± 0.75 as cut-off values for SI_z and SP_z:

- Popular: $LM_z > 0$ and $LL_z < 0$ and $SP_z > +0.75$
- Rejected: $LM_z < 0$ and $LL_z > 0$ and $SP_z < -0.75$
- Neglected: $LM_z < 0$ and $LL_z < 0$ and $SI_z < -0.75$
- Controversial: $LM_z > 0$ and $LL_z > 0$ and $SI_z > +0.75$
- Average: all remaining students

This classification system (hereinafter referred to as CD2) is more liberal than the CD1 method, in such a way that more children are classified as popular, rejected, neglected, and controversial, but fewer children are classified as average.

Method 3 (FW)

The classification procedure by French and Waas (1985) (hereinafter referred to as FW) is comparable to the CD1 method, but does not require the SI_z and SP_z scores and uses ± 0.5 as cut-off values for LL_z and LM_z:

- Popular: $LM_z > 0.5$ and $LL_z < -0.5$
- Rejected: $LM_z < -0.5$ and $LL_z > 0.5$
- Neglected: $LM_z < -0.5$ and $LL_z < -0.5$
- Controversial: $LM_z > 0.5$ and $LL_z > 0.5$
- Average: all remaining students

Method 4 (SVLLS)

Schaughency et al. (1992) recommend assigning the children whose score is above the class median on LM and below the class median on LL to the popular group. Those scoring below the class median on LM and above the class median on LL form the rejected group. Children scoring below the class median on both LM and LL form the neglected group, while those scoring above the class median on both LM and LL form the controversial group. All remaining children form the average group. This classification procedure is hereinafter referred to as SVLLS.

Integrated Teacher Report Form (for Internalizing Behavior) (ITRF-I)

The German version of the ITRF-I was used to assess internalizing behavior in class (Volpe et al., 2020). The screening questionnaire (18 items) was constructed from the subscales of anxious-depressive behavior ($\alpha = 0.87$) and social withdrawal ($\alpha = 0.88$). For each student, class teachers were asked to indicate their level of concern using a four-point scale ranging from "no

---

[1]Please note the debate about the distinction between sociometric and peer-perceived popularity (Parkhurst and Hopmeyer, 1998).

concern" to "strong concern." The anxious-depressive behavior scale consists of 11 items (example: "Acts fearful.") and the social withdrawal scale of 7 items (example: "Does not respond to others' attempts to socialize."). Both narrowband scales can be subsumed into the broadband scale of internalizing behavior.

Four students were excluded due to missing values on more than four items. 253 students had one missing value and 23 students had two to three missing values. Hence, 2047 students (88%) had no missing values. Missing data are imputed by means of predictive mean matching (Eekhout et al., 2014).

Statistical Analysis

To describe the association between internalizing behavior and sociometric status, internalizing behavior is regressed on sociometric status. Four different sociometric status classification methods are applied and both broadband and narrowband internalizing behavior scales are used as criteria (this analysis strategy is schematically illustrated in **Figure 1**), while all regression models are formulated as linear mixed effects regression models (restricted maximum likelihood estimation; nested data structure: students within classes), i.e., random intercept and slope models (Bates et al., 2015).

The sociometric status is used as a dummy variable. The reference is the average sociometric status. Therefore, the intercept of each regression model is interpretable as the expected average internalizing behavior for sociometrically average children. Since all the internalizing behavior scales are standardized ($M = 0$, $SD = 1$), the intercept represents the average internalizing behavior for the average sociometric group as a deviation from the overall sample mean ($M = 0$) in units of standard deviation. The regression parameters (B) for all the other sociometric status groups are interpretable as the difference in internalizing behavior (in units of standard deviation) between the average sociometric group and another sociometric group (popular, rejected, neglected, or controversial). Sex and grade (effect coding: boys vs. girls and grades 1 and 2 vs. grades 3 and 4) changed the regression coefficients for the sociometric status only marginally. Therefore, both variables are not included as covariates (parsimonious model).

In terms of the *New Statistics Approach*, effect sizes and confidence intervals (instead of *p*-values) are reported (Cumming, 2014). All statistical analyses were conducted in R 3.6.0 and with additional R packages (Bates et al., 2015; Barton, 2019; Lüdecke, 2019).

RESULTS

Descriptive Results
Sociometric Status

The number of students classified into the sociometric status groups by the four different sociometric classification methods (CD1, CD2, FW, and SVLLS) is displayed in **Table 1**. It is apparent that the proportion of children assigned to a sociometric group varies across methods [e.g., neglected children: 9% (CD1), 13% (CD2), 4% (FW), and 8% (SVLLS)]. With regard to sociometric neglect, Fleiss' Kappa of 0.67 indicates

a substantial inter-method agreement, albeit 213 (9%) children are simultaneously classified as neglected, popular, rejected, and average, while only 90 (4%) children are congruently classified as neglected. These results underscore the fact that there is no consensus (across sociometric classification methods) about the state of sociometric neglect (Kulawiak and Wilbert, 2019).

Internalizing Behavior

Descriptive parameters of the internalizing behavior scales are displayed in **Table 2**. The correlation between the narrowband scales of internalizing behavior (anxious-depressive behavior and social withdrawal) is moderate ($r = 0.53$), while the broadband scale of internalizing behavior is highly correlated with the narrowband scales (anxious-depressive behavior: $r = 0.91$; social withdrawal: $r = 0.84$). There is some dependence between the variances in the scales of internalizing behavior and the class membership (e.g., social withdrawal: $ICC = 0.12$).

Main Results

Sociometric status is determined by means of four different sociometric classification methods (CD1, CD2, FW, and SVLLS). The broadband and narrowband scales of internalizing behavior are regressed on the different sociometric statuses. The results of the different regression models are displayed in **Table 3**.

Do Different Sociometric Status Classification Methods Give Contradictory Results on Internalizing Behavior in Neglected Students?

In the following, the question is considered whether the application of the different sociometric classification methods leads to contradictory results regarding internalizing behavior in neglected children. The rank order of sociometric groups in terms of the level of internalizing behavior is crucial for judging whether the results of the different sociometric classification methods are contradictory. If all sociometric methods show the same rank order of sociometric groups, the results are consistent. If the rank orders differ, the results are contradictory. If the rank of the neglected group differs across rank orders (i.e., across methods), the results are contradictory with regard to the neglected children. Differences between rank orders are described in more detail (see also **Figure 2**).

Broadband scale of internalizing behavior

Three methods (CD1, CD2, and FW) show the following rank order of sociometric groups in terms of the level of internalizing behavior (ascending): popular, average, controversial, neglected, and rejected (e.g., FW: $B_{popular} = -0.27$, $B_{average} = -0.02$, $B_{controversial} = 0.06$, $B_{neglected} = 0.18$, $B_{rejected} = 0.49$). One method (SVLLS) shows a different rank order (ascending): popular, controversial, average, neglected, and rejected (SVLLS: $B_{popular} = -0.24$, $B_{controversial} = -0.13$, $B_{average} = -0.02$, $B_{neglected} = 0.17$, $B_{rejected} = 0.36$). As both rank orders (i.e., all sociometric methods) show that rejected and neglected children are the most affected and popular children are the least affected by internalizing problems, there are no contradictory results regarding these three sociometric groups. But, in the case of the first rank order (CD1, CD2, and FW), average and controversial

TABLE 1 | Number of students classified into sociometric status groups by different sociometric status classification methods.

Method	Sociometric status					
	Average	Popular	Rejected	Neglected	Controversial	Σ
CD1	1, 274 (55%)	370 (16%)	364 (16%)	211 (9%)	115 (5%)	2, 334 (100%)
CD2	845 (36%)	555 (24%)	462 (20%)	293 (13%)	179 (8%)	2, 334 (100%)
FW	1, 287 (55%)	505 (22%)	402 (17%)	100 (4%)	40 (2%)	2, 334 (100%)
SVLLS	646 (28%)	658 (28%)	640 (27%)	185 (8%)	205 (9%)	2, 334 (100%)
Fleiss' Kappa	0.46	0.74	0.76	0.67	0.54	–
Congruent classification	423 (18%)	333 (14%)	327 (14%)	90 (4%)	32 (1%)	1, 205 (52%)
Multiple classification	1, 129 (48%)	385 (17%)	335 (14%)	213 (9%)	212 (9%)	1, 129 (48%)

CD, Coie and Dodge (1983); FW, French and Waas (1985); SVLLS, Schaughency et al. (1992).

children differ in that controversial children show a higher level of internalizing behavior (e.g., FW: $B_{controversial} = 0.06$), although this difference is very small. The opposite is true in the case of the second rank order (SVLLS), where average and controversial children differ in that controversial children show a lower level of internalizing behavior (SVLLS: $B_{controversial} = -0.13$), although this difference is very small. Hence, there are contradictory results regarding the question of whether controversial or average children are more affected by internalizing problems, although the differences between sociometric groups (controversial vs. average) that indicate this contradiction are very small.

Narrowband scale of anxious-depressive behavior

Two methods (CD1 and CD2) show the following rank order of sociometric groups in terms of the level of anxious-depressive behavior (ascending): popular, average, controversial, neglected, and rejected (e.g., CD1: $B_{popular} = -0.20$, $B_{average} = -0.06$, $B_{controversial} = 0.11$, $B_{neglected} = 0.17$, $B_{rejected} = 0.50$). The FW method shows a different rank order (ascending): popular, average, neglected, controversial, and rejected (FW: $B_{popular} = -0.24$, $B_{average} = -0.01$, $B_{neglected} = 0.01$, $B_{controversial} = 0.16$, $B_{rejected} = 0.41$). In the case of the first rank order (CD1 and CD2), controversial and neglected children differ in that neglected children show a higher level of internalizing behavior (e.g., CD1: $B_{controversial} = 0.11$; $B_{neglected} = 0.17$), although this difference is very small. The opposite is true in the case of the second rank order (FW), where controversial and neglected children differ in that controversial children show a higher level of internalizing behavior (FW: $B_{controversial} = 0.16$; $B_{neglected} = 0.01$), although this difference is very small. Hence,

TABLE 2 | Correlations (Pearson's correlation), means, standard deviations, and intraclass correlations of ITRF-I scales (broadband and narrowband).

Measure	1	2	M	SD	ICC[a]
(1) Internalizing behavior (broadband scale)			5.03	6.70	0.16
(2) Anxious-depressive behavior (narrowband scale)	0.91		2.93	4.33	0.15
(3) Social withdrawal (narrowband scale)	0.84	0.53	2.10	3.30	0.12

All parameters are reported with regard to the raw data. ITRF-I, Integrated Teacher Report Form for Internalizing Behavior. [a]Proportion of variance explained by nested data structure (students within classes).

there are contradictory results regarding neglected children (whether neglected or controversial children are more affected by anxious-depressive behavior), although the differences between sociometric groups (neglected vs. controversial) that indicate this contradiction are very small. A different rank order is also the case with the SVLLS method (ascending): popular, controversial, average, neglected, and rejected (SVLLS: $B_{popular} = -0.21$, $B_{controversial} = -0.02$, $B_{average} = -0.01$, $B_{neglected} = 0.06$, $B_{rejected} = 0.29$). As the neglected and rejected children are at the top of this rank order [as with the first rank order (CD1 and CD2)], this rank order also conflicts with the second rank order (FW), where the controversial and rejected children are at the top. Additionally, there is a contradiction regarding the question of whether average or controversial children are more affected by anxious-depressive behavior. The first rank order (CD1 and CD2) indicates a higher level of anxious-depressive behavior for the controversial group (e.g., CD1: $B_{controversial} = 0.11$), while the third rank order (SVLLS) indicates a higher level of anxious-depressive behavior for the average group (SVLLS: $B_{controversial} = -0.02$), although the differences between sociometric groups (average vs. controversial) that indicate this contradiction are very small.

Narrowband scale of social withdrawal

Three methods (CD1, CD2, and FW) show the following rank order of sociometric groups in terms of the level of social withdrawal behavior (ascending): popular, controversial, average, neglected, and rejected (e.g., FW: $B_{popular} = -0.24$, $B_{controversial} = -0.08$, $B_{average} = -0.02$, $B_{neglected} = 0.38$, $B_{rejected} = 0.45$). The SVLLS method shows a different rank order (ascending): controversial, popular, average, neglected, and rejected (e.g., SVLLS: $B_{controversial} = -0.22$, $B_{popular} = -0.21$, $B_{average} = -0.02$, $B_{neglected} = 0.27$, $B_{rejected} = 0.34$). As both rank orders (i.e., all sociometric methods) show that rejected and neglected children are the most affected by withdrawal behavior and because average children are middle-ranking, there are no contradictory results regarding these three sociometric groups. But, in the case of the first rank order (CD1, CD2, and FW), popular and controversial children differ in that controversial children show a higher level of internalizing behavior (e.g., FW: $B_{popular} = -0.24$, $B_{controversial} = -0.08$), although this difference is very small. The opposite is true in the case of the second rank order (SVLLS), where popular and

TABLE 3 | Internalizing behavior (broadband and narrowband scales) regressed on sociometric status.

	Broadband scale: internalizing behavior (ITRF-I)			Narrowband scale: anxious-depressive behavior (ITRF-I)			Narrowband scale: social withdrawal (ITRF-I)		
	B	**95% CI**	**mR²/cR²**	**B**	**95% CI**	**mR²/cR²**	**B**	**95% CI**	**mR²/cR²**
CD1			0.06/0.28			0.04/0.24			0.06/0.27
Intercept[a]	−0.07	−0.16 – 0.01		−0.06	−0.14 – 0.03		−0.07	−0.15 – 0.01	
Popular	−0.22	−0.33 – −0.12		−0.20	−0.31 – −0.10		−0.19	−0.29 – −0.09	
Rejected	0.61	0.47 – 0.75		0.50	0.37 – 0.64		0.57	0.42 – 0.72	
Neglected	0.30	0.16 – 0.45		0.17	0.03 – 0.32		0.40	0.22 – 0.57	
Controversial	0.02	−0.15 – 0.19		0.11	−0.07 – 0.29		−0.13	−0.30 – 0.04	
CD2			0.06/0.27			0.04/0.23			0.07/0.28
Intercept[a]	−0.07	−0.16 – 0.03		−0.04	−0.13 – 0.06		−0.09	−0.17 – 0.00	
Popular	−0.21	−0.30 – −0.11		−0.19	−0.29 – −0.09		−0.17	−0.26 – −0.07	
Rejected	0.53	0.40 – 0.65		0.40	0.28 – 0.51		0.55	0.41 – 0.68	
Neglected	0.25	0.13 – 0.38		0.11	−0.01 – 0.24		0.38	0.23 – 0.54	
Controversial	−0.02	−0.18 – 0.13		0.06	−0.11 – 0.22		−0.11	−0.26 – 0.03	
FW			0.06/0.26			0.04/0.23			0.05/0.23
Intercept[a]	−0.02	−0.11 – 0.07		−0.01	−0.10 – 0.08		−0.02	−0.10 – 0.06	
Popular	−0.27	−0.37 – −0.18		−0.24	−0.34 – −0.15		−0.24	−0.33 – −0.14	
Rejected	0.49	0.37 – 0.60		0.41	0.29 – 0.53		0.45	0.33 – 0.58	
Neglected	0.18	−0.03 – 0.40		0.01	−0.20 – 0.21		0.38	0.15 – 0.61	
Controversial	0.06	−0.26 – 0.38		0.16	−0.19 – 0.50		−0.08	−0.41 – 0.24	
SVLLS			0.05/0.23			0.04/0.20			0.05/0.22
Intercept[a]	−0.02	−0.12 – 0.09		−0.01	−0.12 – 0.09		−0.02	−0.12 – 0.08	
Popular	−0.24	−0.34 – −0.14		−0.21	−0.31 – −0.11		−0.21	−0.32 – −0.11	
Rejected	0.36	0.26 – 0.47		0.29	0.19 – 0.39		0.34	0.22 – 0.46	
Neglected	0.17	0.01 – 0.32		0.06	−0.09 – 0.21		0.27	0.10 – 0.44	
Controversial	−0.13	−0.27 – 0.02		−0.02	−0.17 – 0.13		−0.22	−0.37 – −0.07	

Linear mixed effects regression models (restricted maximum likelihood estimation; random intercept and slope models; nested data structure: students within classes). Internalizing behavior scales (broadband and narrowband) are standardized (M = 0, SD = 1). [a]The sociometric status is dummy coded (reference: average); mR²/cR², marginal R²/conditional R²; ITRF-I, Integrated Teacher Report Form for Internalizing Behavior; CD, Coie and Dodge (1983); FW, French and Waas (1985); SVLLS: Schaughency et al. (1992).

controversial children differ in that controversial children show a lower level of internalizing behavior (SVLLS: $B_{popular} = -0.21$, $B_{controversial} = -0.22$), although this difference is very small. Hence, there are contradictory results regarding the question of whether popular or controversial children are more affected by withdrawal behavior, although the differences between sociometric groups (popular vs. controversial) that indicate this contradiction are very small.

Do Narrowband Behavior Scales Show Differentiated Effect Sizes and Does the Broadband Behavior Scale Underestimate the Association Between Internalizing Behavior and Sociometric Neglect?

The question is considered here whether the application of narrowband scales shows differentiated effect sizes and whether the use of the broadband scale underestimates the association (i.e., the narrowband scales show stronger effect sizes).

Differences in the prediction between the narrowband scales (differentiated effects) are apparent, as in all analyses the narrowband scale of social withdrawal is more strongly associated with sociometric neglect (e.g., CD2: $B_{neglected} = 0.38$) than the narrowband scale of anxious-depressive behavior (e.g., CD2: $B_{neglected} = 0.11$) (the effect sizes differ many times over). An underestimation of the association by the broadband scale is evident in that, in all analyses, the narrowband scale of

social withdrawal is more strongly associated with sociometric neglect (e.g., FW: $B_{neglected} = 0.38$) than the broadband scale of internalizing behavior (e.g., FW: $B_{neglected} = 0.18$).

Differences in the prediction between the narrowband scales are also noticeable with regard to the controversial sociometric status, as the narrowband scale of social withdrawal is negatively associated with sociometric controversiality (e.g., CD1: $B_{controversial} = -0.13$) while the narrowband scale of anxious-depressive behavior is positively associated (e.g., CD1: $B_{controversial} = 0.11$). This differentiated effect (positive and negative association) is not visible when using the broadband scale of internalizing behavior, i.e., the effect size is close to zero (e.g., CD1: $B_{controversial} = 0.02$).

DISCUSSION

In this study, we examined the question of whether and to what extent the conceptualization and measurement of the social inclusion of students with SEN is relevant for the generation of consistent research results. This issue remains important because poor measurement of social inclusion has been linked to equivocal research results on the social inclusion of students with SEN in mainstream schools (Chambers and Kay, 1992; Avramidis et al., 2017). Therefore, in the present paper, issues
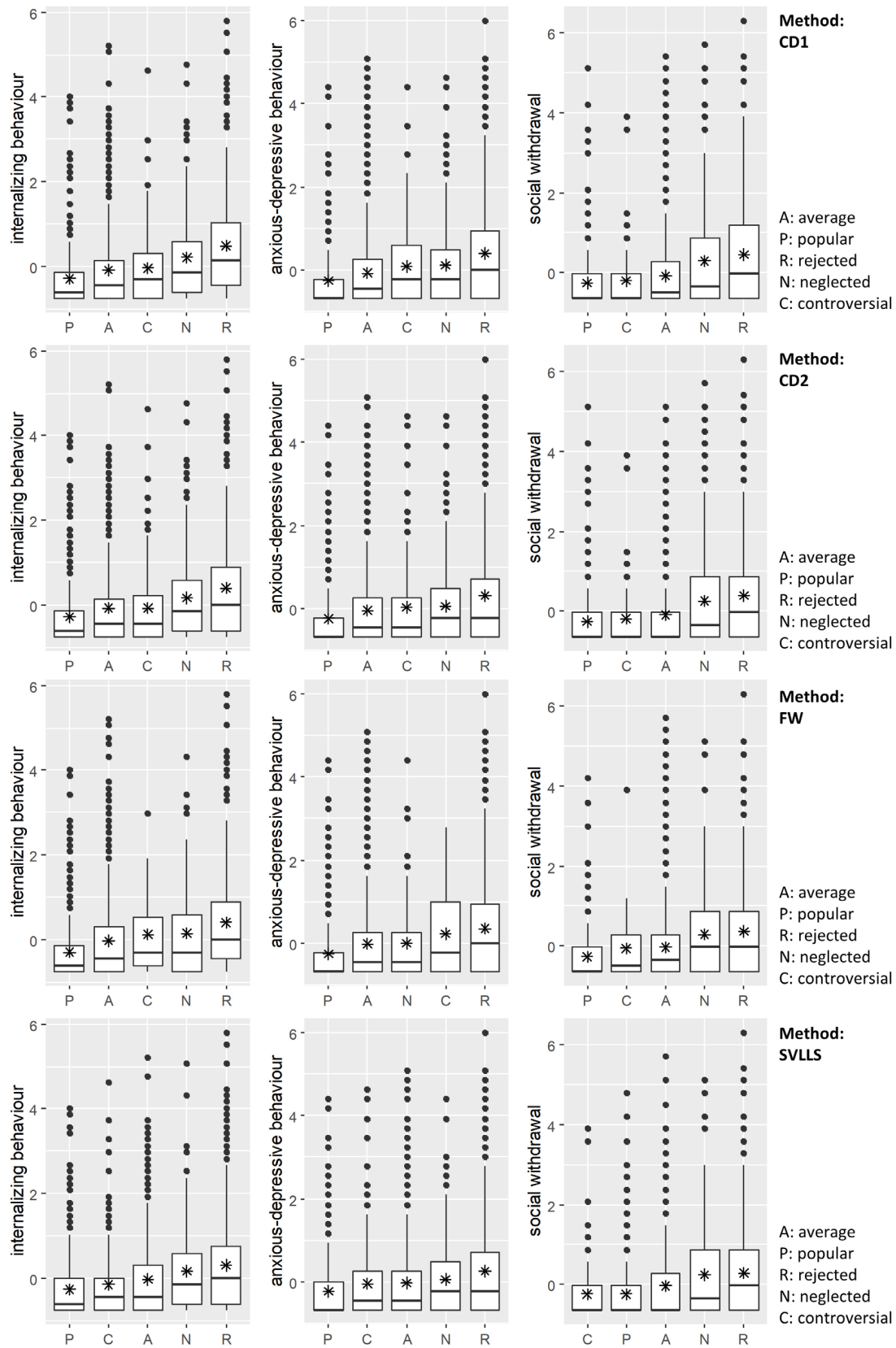
FIGURE 2 | Broadband scale of internalizing behavior and narrowband scales of internalizing behavior (anxious-depressive behavior and social withdrawal) by sociometric groups (boxplots in ascending rank order).

of a specific assessment approach of social inclusion were examined, namely the sociometric status classification, which is widely used in research on social inclusion of students with SEN. At the same time, we outlined the equivocal evidence on the internalizing behavior of sociometrically neglected children. Internalizing problems in children belong to the category of SEN called emotional and behavioral difficulties. An undifferentiated (broadband) perspective on children with SEN has also been linked to contradictory research results (O'Mara et al., 2012). Therefore, in the present paper, the operationalization of internalizing behavior was likewise examined, namely the operationalization of broadband and narrowband dimensions of behavior.

The analyses and results presented here provide initial indications that methodological approaches can lead to conflicting results on the internalizing behavior in sociometrically neglected children (but also with regard to the other sociometric groups), although the differences in behavior between sociometric groups that indicate these contradictions are small. The contradictory results are not only due to the application of different sociometric classification methods, but also to different operationalizations of internalizing behavior (narrowband and broadband scales).

Contradictions regarding the question of whether neglected children are more affected by anxious-depressive behavior than average and controversial children are due to the application of different sociometric classification methods. These contradictory results reflect the current equivocal state of research. For example, one study implies that neglected children are the most affected and controversial children are the least affected by anxiety (La Greca et al., 1988), while another study indicates higher anxiety among controversial children (Crick and Ladd, 1993). The results are conflicting and both studies use different sociometric classification methods.

The two narrowband scales show differentiated effects, as the narrowband scale of social withdrawal is more strongly associated with sociometric neglect than the narrowband scale of anxious-depressive behavior (the effect sizes differ many times over). At the same time, an underestimation of the association by the broadband scale is evident in that the narrowband scale of social withdrawal is more strongly associated with sociometric neglect than the broadband scale of internalizing behavior. Hence, the differentiation between narrow facets of internalizing behavior and the application of narrowband scales provide a deeper understanding of internalizing behavior in neglected children. These results support the assumption that there is no benefit (but rather harm) in summarizing distinctive behavior problems (operationalized by narrowband scales) into one broad category of behavior (operationalized by a broadband scale) (Tandon et al., 2009; Kulawiak et al., under review).

Limitations

Critically, it should be noted that the narrowband scale of anxious-depressive behavior used in this study summarizes two distinctive behavioral problems, namely anxiety and depression. Strictly speaking, a more differentiated view of anxiety and depression (in the sense of two different narrowband scales) could provide an even more detailed insight into neglected children's internalizing behavior. Beyond that, this study covers just some dimensions and not the entire spectrum of dimensions of internalizing behavior. This, in turn, may lead to the premature conclusion that internalizing problems are not related to the neglected sociometric status, as other relevant narrowband dimensions of internalizing behavior may not have been evaluated.

Conclusion

The systematic investigation presented here provides initial indications that methodological approaches can lead to conflicting results. Earlier contradictory evidence on the internalizing behavior of neglected students must therefore be seen in a different light: the reasons for previously conflicting results may actually be methodological. The main methodological issue is that even after decades of sociometric research there is still no consensus method for identifying neglected children. This reflects the lack of a sufficient conceptualization of sociometric neglect. Early on, a clearer definition of sociometric neglect was demanded (Gottman, 1977). However, even decades later, the conceptualization of sociometric neglect is still in progress (Kulawiak and Wilbert, 2019) and new sociometric classification methods appear at regular intervals (García Bacete and Cillessen, 2017), some of which emphasize the need for a more accurate representation in the sense of a continuous metric of the sociometric status (DeRosier and Thomas, 2003; Kulawiak and Wilbert, 2019). Which sociometric classification method provides the most reliable research results? Given the arbitrariness of the sociometric classification rules, this question is hard to answer. In order to improve the operationalization of sociometric neglect, it is necessary to clarify how neglect affects internalizing behavior (or vice versa). For example, there is a difference between passive and active neglect (others ignore vs. avoid the individual) (Leary, 1990). This kind of nuanced information is not presented by current sociometric classification methods. Both types of neglect can have different effects on a child's internalizing behavior. In the context of sociometric research on internalizing behavior, therefore, a more detailed conceptualization of sociometric neglect is needed. Measuring approaches other than sociometry (e.g., observation of peer interactions or self-reports of feelings) could be an enrichment for the assessment of neglect, as these procedures can capture more nuanced information (passive and active neglect or subjective feelings of being neglected). Besides that, many of the sociometric studies on internalizing behavior are cross-sectional (including the present study) and therefore do not investigate the causal relationship or causal direction between internalizing behavior and sociometric neglect. Hence, priority must be given to a theoretical model that postulates a causal link between sociometric neglect and internalizing behavior (or vice versa). For example, the experience of being neglected (or ignored) by peers may lead to social-evaluative concerns that may in turn prompt the inhibition of sociable behavior (Asendorpf, 1990), i.e., children would act socially withdrawn. But, the experience of being neglected may also trigger other internalizing reactions (Leary, 1990), e.g., the experience of being neglected may reinforce fears of future social interactions (social anxiety). In order to derive

a plausible causal model, the relevant narrowband dimensions of internalizing behavior must be theory-driven and clearly identified (instead of summarizing of all internalizing problems into one broadband category).

DATA AVAILABILITY STATEMENT

The datasets generated for this study will not be made publicly available due to data protection regulation.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

REFERENCES

Achenbach, T. M., Ivanova, M. Y., Rescorla, L. A., Turner, L. V., and Althoff, R. R. (2016). Internalizing/externalizing problems: review and recommendations for clinical and research applications. *J. Am. Acad. Child Adolesc. Psychiatry* 55, 647–656. doi: 10.1016/j.jaac.2016.05.012

Asendorpf, J. B. (1990). Development of inhibition during childhood: evidence for situational specificity and a two-factor model. *Dev. Psychol.* 26, 721–730. doi: 10.1037/0012-1649.26.5.721

Avramidis, E., Avgeri, G., and Strogilos, V. (2018). Social participation and friendship quality of students with special educational needs in regular Greek primary schools. *Eur. J. Spec. Needs Educ.* 33, 221–234. doi: 10.1080/08856257. 2018.1424779

Avramidis, E., Strogilos, V., Aroni, K., and Kantaraki, C. T. (2017). Using sociometric techniques to assess the social impacts of inclusion: some methodological considerations. *Educ. Res. Rev.* 20, 68–80. doi: 10.1016/j.edurev. 2016.11.004

Barton, K. (2019). *MuMIn: Multi-Model Inference. R package Version 1.43.6.* Available online at: https://cran.r-project.org/web/packages/MuMIn/MuMIn. pdf (accessed March 23, 2020).

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bierman, K. L. (1987). The clinical significance and assessment of poor peer relations: peer neglect versus peer rejection. *J. Dev. Behav. Pediatr. JDBP* 8, 233–240.

Boivin, M., Poulin, F., and Vitaro, F. (1994). Depressed mood and peer rejection in childhood. *Dev. Psychopathol.* 6, 483. doi: 10.1017/S0954579400006064

Bukowski, W. M., Cillessen, A. H. N., and Velásquez, A. M. (2012). "Peer Ratings," in *Handbook of Developmental Research Methods*, eds B. Paul, Laursen, T. D. Little, and N. A. Card (New York, NY: Guilford Press), 211–228.

Burton, C. B., and Krantz, M. (1990). Predicting adjustment in middle childhood from early peer status. *Early Child Dev. Care* 60, 89–100. doi: 10.1080/ 0300443900600108

Cantrell, V. L., and Prinz, R. J. (1985). Multiple perspectives of rejected, neglected, and accepted children: Relation between sociometric status and behavioral characteristics. *J. Consult. Clin. Psychol.* 53, 884–889. doi: 10.1037//0022-006X. 53.6.884

Chambers, S. M., and Kay, R. W. (1992). Research on Social Integration: what are the problems? *Int. J. Disabil. Dev. Educ.* 39, 47–59. doi: 10.1080/ 0156655920390106

Coie, J. D., and Dodge, K. A. (1983). Continuities and changes in children's social status: a five-year longitudinal study. *Merrill-Palmer Q.* 29, 261–282.

Coie, J. D., and Kupersmidt, J. B. (1983). A behavioral analysis of emerging social status in boys'. *Groups. Child Dev* 54, 1400–1416. doi: 10.1111/j.1467-8624. 1983.tb00056.x

Crews, S. D., Bender, H., Cook, C. R., Gresham, F. M., Kern, L., and Vanderwood, M. (2007). Risk and protective factors of emotional and/or behavioral disorders

in children and adolescents: a mega-analytic synthesis. *Behav. Disord.* 32, 64–77. doi: 10.1177/019874290703200201

Crick, N. R., and Ladd, G. W. (1993). Children's perceptions of their peer experiences: attributions, loneliness, social anxiety, and social avoidance. *Dev. Psychol.* 29, 244–254. doi: 10.1037/0012-1649.29.2.244

Cumming, G. (2014). The new statistics: why and how. *Psychol. Sci.* 25, 7–29. doi: 10.1177/0956797613504966

Delgado, B., Inglés, C. J., Aparisi, D., and García-Fernández, J. M. (2016). Social anxiety and sociometric nomination in spanish students of compulsory secondary education. *Span. J. Psychol.* 19, E41. doi: 10.1017/sjp.2016.47

DeRosier, M. E., and Thomas, J. M. (2003). Strengthening sociometric prediction: scientific advances in the assessment of children's peer relations. *Child Dev* 74, 1379–1392. doi: 10.1111/1467-8624.00613

Eekhout, I., de Vet, H. C. W., Twisk, J. W. R., Brand, J. P. L., de Boer, M. R., and Heymans, M. W. (2014). Missing data in a multi-item instrument were best handled by multiple imputation at the item score level. *J. Clin. Epidemiol.* 67, 335–342. doi: 10.1016/j.jclinepi.2013.09.009

Epkins, C. C., and Heckler, D. R. (2011). Integrating etiological models of social anxiety and depression in youth: evidence for a cumulative interpersonal risk model. *Clin. Child Fam. Psychol. Rev.* 14, 329–376. doi: 10.1007/s10567-011-0101-8

Frederickson, N., Simmonds, E., Evans, L., and Soulsby, C. (2007). Assessing the social and affective outcomes of inclusion. *Br. J. Spec. Educ.* 34, 105–115. doi: 10.1111/j.1467-8578.2007.00463.x

French, D. C., and Waas, G. A. (1985). Behavior problems of peer-neglected and peer-rejected elementary-age children: parent and teacher perspectives. *Child Dev* 56, 246–252.

Friedman, A. H. (2004). *Specific Aspects of Social Anxiety Related to Children's Peer Sociometric Status.* Morgantown, WV: West Virginia University.

Frostad, P., Mjaavatn, P. E., and Pijl, S. J. (2011). The stability of social relations among adolescents with special educational needs (SEN) in regular schools in Norway. *Lond. Rev. Educ.* 9, 83–94. doi: 10.1080/14748460.2011. 550438

García Bacete, F. J., and Cillessen, A. H. N. (2017). An adjusted probability method for the identification of sociometric status in classrooms. *Front. Psychol.* 8:1836. doi: 10.3389/fpsyg.2017.01836

Garrote, A., Sermier Dessemontet, R., and Moser Opitz, E. (2017). Facilitating the social participation of pupils with special educational needs in mainstream schools: a review of school-based interventions. *Educ. Res. Rev.* 20, 12–23. doi: 10.1016/j.edurev.2016.11.001

Garrote, G. (2017). Relationship between the social participation and social skills of pupils with an intellectual disability: a study in inclusive classrooms. *Frontline Learn. Res.* 5:1–15. doi: 10.14786/flr.v5i1.266

Gerullis, A., and Huber, C. (2018). I agree to sit next to you. Does that mean I like you? Measuring using the wrong tapeline - the lack of 'Social Distance' measures for inclusive school development and research - a review of the literature. *J. Res. Spec. Educ. Needs* 18, 124–135. doi: 10.1111/1471-3802.12399

Goodman, A., Lamping, D. L., and Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the strengths and difficulties questionnaire (SDQ): data from british parents, teachers and children. *J. Abnorm. Child Psychol.* 38, 1179–1191. doi: 10.1007/s10802-010-9434-x

Gottman, J. M. (1977). Toward a definition of social isolation in children. *Child Dev.* 48, 513. doi: 10.2307/1128647

Hecht, D. B., Inderbitzen, H. M., and Bukowski, A. L. (1998). The relationship between peer status and depressive symptoms in children and adolescents. *J. Abnorm. Child Psychol.* 26, 153–160.

Henke, T., Bogda, K., Lambrecht, J., Bosse, S., Koch, H., Maaz, K., et al. (2017). Will you be my friend? A multilevel network analysis of friendships of students with and without special educational needs backgrounds in inclusive classrooms. *Z. Für Erzieh.* 20, 449–474. doi: 10.1007/s11618-017-0767-x

Hennemann, T., Wilbert, J., Krull, J., and Urton, K. (2018). *Schools On Their Journey Towards Inclusion: Mettmann 2.0.* Available online at: https://www.hf.uni-koeln.de/38801 (accessed March 23, 2020).

Howe, C. (2010). "Social and Personal Adjustment," in *Peer Groups and Children's Development Understanding Children's Worlds* (Malden, MA: Wiley-Blackwell), 142–163.

Hubbard, J. A. (2001). Emotion expression processes in children's peer interaction: the role of peer rejection, aggression, and gender. *Child Dev.* 72, 1426–1438. doi: 10.1111/1467-8624.00357

Hymel, S., and Rubin, K. H. (1985). "Children with peer relationships and social skills problems: Conceptual, methodological, and developmental issues," in *Annals of Child Development: a Research Annual*, Vol. 2, ed. G. J. Whitehurst (London: Jai Press), 251–297.

Kaya, A. (2007). Sociometric status, depression, and locus of control among turkish early adolescents. . *Soc. Behav. Personal. Int. J.* 35, 1405–1414. doi: 10.2224/sbp.2007.35.10.1405

Kershaw, P., and Sonuga-Barke, E. (1998). Emotional and behavioural difficulties: is this a useful category? The implications of clustering and co-morbidity - the relevance of a taxonomic approach. *Educ. Child Psychol.* 15, 45–55.

Kingery, J. N., Erdley, C. A., Marshall, K. C., Whitaker, K. G., and Reuter, T. R. (2010). Peer experiences of anxious and socially withdrawn youth: an integrative review of the developmental and clinical literature. *Clin. Child Fam. Psychol. Rev.* 13, 91–128. doi: 10.1007/s10567-009-0063-2

Koster, M., Nakken, H., Pijl, S. J., and van Houten, E. (2009). Being part of the peer group: a literature study focusing on the social dimension of inclusion in education. *Int. J. Incl. Educ.* 13, 117–140. doi: 10.1080/13603110701284680

Krull, J., Wilbert, J., and Hennemann, T. (2014). The social and emotional situation of first graders with classroom behavior problems and classroom learning difficulties in inclusive classes. *Learn. Disabil. Contemp. J.* 12, 169–190.

Kulawiak, P. R., and Wilbert, J. (2019). Introduction of a new method for representing the sociometric status within the peer group: the example of sociometrically neglected children. *Int. J. Res. Method Educ.* 1–19. doi: 10.1080/1743727X.2019.1621830

Kupersmidt, J. B., and Patterson, C. J. (1991). Childhood peer rejection, aggression, withdrawal, and perceived competence as predictors of self-reported behavior problems in preadolescence. *J. Abnorm. Child Psychol.* 19, 427–449. doi: 10.1007/BF00919087

La Greca, A. M., Dandes, S. K., Wick, P., Shaw, K., and Stone, W. L. (1988). Development of the social anxiety scale for children: reliability and concurrent validity. *J. Clin. Child Psychol.* 17, 84–91. doi: 10.1207/s15374424jccp1701_11

Leary, M. R. (1990). Responses to social exclusion: social anxiety, jealousy, loneliness, depression, and low Self-Esteem. *J. Soc. Clin. Psychol.* 9, 221–229. doi: 10.1521/jscp.1990.9.2.221

Lüdecke, D. (2019). *sjstats: Collection of Convenient Functions for Common Statistical Computations. R package version 0.17.6.* Available online at: https://cran.r-project.org/web/packages/sjstats/sjstats.pdf (accessed March 23, 2020).

Marks, P. E. L., Babcock, B., Cillessen, A. H. N., and Crick, N. R. (2013). The effects of participation rate on the internal reliability of peer nomination measures: participation and sociometric reliability. *Soc. Dev.* 22, 609–622. doi: 10.1111/j.1467-9507.2012.00661.x

Mayeux, L., Bellmore, A. D., and Cillessen, A. H. N. (2007). Predicting changes in adjustment using repeated measures of sociometric status. *J. Genet. Psychol.* 168, 401–424. doi: 10.3200/GNTP.168.4.401-424

McMullen, J. A., Veermans, K., and Laine, K. (2014). Tools for the classroom? An examination of existing sociometric methods for teacher use. *Scand. J. Educ. Res.* 58, 624–638. doi: 10.1080/00313831.2013.838694

Newcomb, A. F., Bukowski, W. M., and Pattee, L. (1993). Children's peer relations: a meta-analytic review of popular, rejected, neglected, controversial, and average sociometric status. *Psychol. Bull.* 113, 99–128. doi: 10.1037/0033-2909.113.1.99

Nakken, H., and Pijl, S. J. (2002). Getting along with classmates in regular schools: a review of the effects of integration on the development of social relationships. *Int. J. Incl. Educ.* 6, 47–61. doi: 10.1080/13603110110051386

O'Mara, A., Akre, B., Munton, T., Marrero-Guillamon, I., Martin, A., Gibson, K., et al. (2012). "Review of the International Literature", in *Curriculum and Curriculum Access Issues for Students with Special Educational Needs in Post-primary Settings: An international review NCSE RESEARCH REPORTS. (Trim)*. 36–81. Available online at: http://ncse.ie/wp-content/uploads/2014/10/Access_to_the_Curriculum_19_10_12-2.pdf (accessed March 23, 2020).

Parkhurst, J. T., and Hopmeyer, A. (1998). Sociometric popularity and peer-perceived popularity: two distinct dimensions of peer status. *J. Early Adolesc.* 18, 125–144. doi: 10.1177/0272431698018002001

Rubin, K. H., Hymel, S., Lemare, L., and Rowden, L. (1989). Children experiencing social difficulties: Sociometric neglect reconsidered. *Can. J. Behav. Sci. Rev. Can. Sci. Comport.* 21, 94–111. doi: 10.1037/h0079775

Rytioja, M., Lappalainen, K., and Savolainen, H. (2019).Behavioural and emotional strengths of sociometrically popular, rejected, controversial, neglected, and average children. *Eur. J. Spec. Needs Educ.* 34, 557–571. doi: 10.1080/08856257.2018.1560607

Schaughency, E. A., Vannatta, K., Langhinrichsen, J., Lally, C., and Seeley, J. (1992). Correlates of sociometric status in school children in Buenos Aires. *J. Abnorm. Child Psychol.* 20, 317–326. doi: 10.1007/BF00916695

Strauss, C. C., Lahey, B. B., Frick, P., Frame, C. L., and Hynd, G. W. (1988). Peer social status of children with anxiety disorders. *J. Consult. Clin. Psychol.* 56, 137–141.

Tandon, M., Cardeli, E., and Luby, J. (2009). Internalizing disorders in early childhood: a review of depressive and anxiety disorders. *Child Adolesc. Psychiatr. Clin. N. Am.* 18, 593–610. doi: 10.1016/j.chc.2009.03.004

Tani, F., and Schneider, B. (1997). Self-reported symptomatology of socially rejected and neglected Italian elementary-school children. *Child Study J.* 27, 301–317.

Terry, R., and Coie, J. D. (1991). A comparison of methods for defining sociometric status among children. *Dev. Psychol.* 27, 867–880. doi: 10.1037/0012-1649.27.5.867

Volpe, R. J., Yeung, J., Casale, G., Krull, J., Briesch, A. M., and Hennemann, T. (2020). Evaluation of a German language school-based universal screening for student social, emotional, and behavioral risk. *Int. J. Sch. Educ. Psychol.* 1–11. doi: 10.1080/21683603.2020.1717699

Whitcomb, S. A. (2018). "Assessing Internalizing Problems," in *Behavioral, Social, and Emotional Assessment of Children and Adolescents*, ed. K. M. Merrell (New York, NY: Routledge), 305–342.

Wilmshurst, L. (2017). *Abnormal Child and Adolescent Psychology: a Developmental Perspective*, 2nd Edn. New York, NY: Routledge.

Woodhouse, S. S., Dykas, M. J., and Cassidy, J. (2012). Loneliness and peer relations in adolescence: loneliness in adolescence. *Soc. Dev.* 21, 273–293. doi: 10.1111/j.1467-9507.2011.0