



# Student-Reported Classroom Climate Pre and Post Teacher Training in Restorative Practices

Constanze Weber\*, Merle Rehder and Leen Vereenooghe

Psychological Intervention in Inclusive Settings, Department of Psychology, Faculty of Psychology and Sports Science, Bielefeld University, Bielefeld, Germany

## OPEN ACCESS

### Edited by:

Joseph Mintz,  
University College London,  
United Kingdom

### Reviewed by:

Xiao Qu,  
University of the West of Scotland,  
United Kingdom  
Caroline Hilton,  
University College London,  
United Kingdom

### \*Correspondence:

Constanze Weber  
constanze.weber@uni-bielefeld.de

### Specialty section:

This article was submitted to  
Special Educational Needs,  
a section of the journal  
Frontiers in Education

Received: 02 June 2021

Accepted: 23 July 2021

Published: 19 August 2021

### Citation:

Weber C, Rehder M and  
Vereenooghe L (2021) Student-  
Reported Classroom Climate Pre and  
Post Teacher Training in  
Restorative Practices.  
Front. Educ. 6:719357.  
doi: 10.3389/feduc.2021.719357

Restorative practices (RP) offer a means to establish positive and caring relationships and could thus foster the mental and scholastic development of students by improving classroom climate. This could benefit both students with and without special educational needs and disabilities (SEND), yet to date no studies evaluated these practices in inclusive educational settings. Here we report the findings of two consecutive studies: a pilot single-group pre-post (Study 1) and a non-randomised controlled study of RP training vs no-intervention control condition (Study 2). Across both studies, 531 students (46.5% female) with a mean age of 11.43 years ( $SD = 1.27$ ) enrolled in the study at pre-test, of which 13.9% had a confirmed diagnosis of SEND and a further 5.7% were considered by teachers to likely have SEND. School and classroom climate, as well as victimisation experiences, emotional well-being and social inclusion of students were assessed using self-report questionnaires. Easy enrolment of schools and students at pre-test indicated that studies investigating the effects of RP training could be feasible. However, in part due to COVID-19 related school closures, student attrition rates of 90 and 77% were observed, for Study 1 and Study 2 respectively. In spite of observed improvements in classroom climate for the intervention group in Study 2, statistical analyses yielded no significant effects of the intervention and there were no moderation effects of students' perceived inclusion and victimisation experiences. Together, these studies provide the first quantitative student data on implementing RP in an inclusive educational setting. We discuss our findings in light of the need for ideas on how to reduce attrition and also consider longer school-wide and single-class implementations of RP.

**Keywords:** restorative practices, inclusive education, classroom climate, relationships in class, german educational context

## INTRODUCTION

School and classroom climate play a prominent role in the academic and psychological development of students (Wang and Degol, 2016; Grewe, 2017). School climate, for example, has an important influence on the psychological development of children and young people (e.g., Koth et al., 2008; Cohen et al., 2009; Schulte-Körne, 2016). Likewise, a positive classroom climate is particularly important for both the school performance (Flook et al., 2005; Eder, 2018) and the psychological development of children and young people (e.g., Grewe, 2003).

Whilst sharing some characteristics, school and classroom climate are two distinct, yet interdependent, multidimensional constructs. School climate refers to the shared patterns of

experience of all people in school life and thus reflects the norms, values, objectives, and the general shaping of interpersonal relationships, teaching and learning practices, and organisational structures (Thapa et al., 2013). Meanwhile, across the varying definitions and aspects of classroom climate there is a consensus that it refers to the socially shared subjective representation of important characteristics of the school class as a learning environment (Eder, 2002). An important aspect of classroom climate involves the relationships between individuals: both among students and between students and their teachers.

Irrespective of the focus on the school or classroom level, a central aspect of the concept of climate is that it is a “collective” construct formed from individual and socially shared perceptions (Eder, 2002). Consequently, climate is considered to be a dynamic rather than a static construct (Wang and Degol, 2016), with some studies suggesting that perceptions of school climate decline during middle and high school (Way et al., 2007; Wang and Dishion, 2012). Meanwhile, for individual students’ class climate perceptions, the student-student relationship and student-teacher relationship appear to be particularly relevant (Eder, 1996).

A positive school and classroom climate benefits both the academic performance and the (social) well-being of all students (Cohen et al., 2009). Hence, installing a positive climate is of particular relevance for inclusive education, as this could improve the academic and psychological outcomes of students with and without special educational needs and disabilities (SEND).

In Germany, inclusive education is a requirement since ratifying the UN Convention on the Rights of Persons with Disabilities (United Nations, 2006) in 2009. In this study, we follow the definition of inclusive education according to the Federal State North Rhine-Westphalia (Lütje-Klose et al., 2017) in that students with SEND are taught together with their peers without SEND. In national empirical educational research, this is probably the most commonly studied concept of inclusion (Grosche, 2015). However, the proportions of students with SEND and the extent of inclusion differ significantly between the federal states in Germany, who implement different inclusion practices and SEND classifications (Heisig, 2018). While the inclusion rate in North Rhine-Westphalia in 2015 was 46.9% for elementary schools (grade 1–4) and 29.9% for secondary schools (lower secondary: grades 5–8, upper secondary: grades 9–12/13), many secondary schools have yet to take steps towards inclusive education (Klemm, 2015). For secondary schools in this federal state, where the current studies were conducted, there are two dominant approaches to the concept of inclusive education. In one approach, one or more inclusive classes are designated per grade, and hence inclusive education is not necessarily offered or strived for in all classes. In the other approach, all classes of each grade are open to SEND students. Irrespective of which classes are designated as inclusive classes, teachers are supported by special needs educators or socio-educational assistants.

Turning to previous studies investigating inclusive educational processes in primary and secondary education, it becomes apparent that there has been more attention for the potential

effects of inclusion on students’ performance (e.g., Cambra and Silvestre, 2003; Huber and Wilbert, 2012) than for its potential effects on classroom climate or on the social participation of SEND students (Crede et al., 2019). The importance of the latter cannot be understated, as students with SEND run the risk of being less accepted by their peers (Koster et al., 2010; Pijl and Frostad, 2010), having less friends (Frostad and Pijl, 2007; Koster et al., 2010; Avramidis, 2013) and are likely to experience the classroom climate more negatively than their classmates without SEND (Koster et al., 2009). These findings are not consistently demonstrated (Spörer et al., 2015; Garrote et al., 2017), however, and the perception of social relations in the classroom and its climate may be influenced by multiple factors (e.g. teacher attitude, type of SEND, inclusion concept).

In response to the student heterogeneity in inclusive classrooms and its challenges regarding the strengthening of social relationships and classroom climate, we therefore investigate a relationship-orientated approach to acknowledge and promote inclusive classroom relationships, namely restorative practices (RP). RP offer an approach to foster positive relationships within the school setting and the larger community, while resolving conflicts constructively (Hendry, 2009), and have been increasingly used in school settings around the world. RP are more common in the judicial context, where they are used as an alternative way of responding constructively to conflict, or discipline and behaviour issues. As these issues also arise in a school setting, the transfer of RP to a school setting is possible (Anfara et al., 2013; Green et al., 2019). The key objectives of RP are context independent and are concerned with community building, improving relationships, and problem solving to settle conflict, while also holding individuals accountable for their behaviour (Hendry, 2009). To that extent, RP comprises a continuum of practices that range from prevention (before an infraction) to intervention (after an infraction). For an overview of the specific restorative methods, see **Supplementary File S1**. The concept of RP in schools thus includes improving relationships not only between and among students, but also with teachers, schools, and entire communities (Anfara et al., 2013), making it a whole-school approach.

Research over the last decade suggests that the use of RP in schools might reduce bullying and improve student-teacher relationships, whereas evidence for its positive effects on school and classroom climate are inconsistent (Weber and Vereenoghe, 2020). Where qualitative studies report improvements (Mirsky, 2007; Costello et al., 2009), these improvements were not consistently confirmed in quantitative studies (e.g., Augustine et al., 2018). For example, Acosta et al. (2019) conducted a cluster randomised controlled trial (RCT) and have found their RP intervention did not improve student ratings of school climate. However, students’ self-reported experience with RP, assessed by asking how often their teachers used specific methods of RP, was significantly associated with improved school climate at post-test. Assuming RP can positively impact the whole school climate and promote positive youth development, it could be a promising approach to improve student cohesion and school

and classroom climate in inclusive schools where students with SEND are more likely to experience lower satisfaction in these areas.

To date, there are no studies investigating the effects of RP in the German school system, nor with specific reference to the challenges of inclusive education. Within the various aspects of school and classroom climate that could be taken into account, the present research focuses primarily on the interpersonal aspects of school and classroom climate, because they are strongly related to both the academic and the psychological development of students (Flook et al., 2005).

Here, we present the findings of two separate studies: a single-group pre-post pilot study (Study 1) and a non-randomised controlled main study (Study 2) in which teachers of inclusive classes received training in RP. Whilst both studies adopted the same outcome measures to assess intervention effects at the student-level, parts of the RP training in the intervention group were revised following the pilot study. These changes are clarified in the respective study descriptions. However, as the implementation and research of RP in the German school system is new, it was not desirable to consider all confounding variables in the study design and we mainly strived to examine the feasibility of our training and study designs. In this regard, RP was not implemented as a whole-school approach. Furthermore, with no RP training currently available in Germany, the research team first trained themselves prior to developing a new training manual.

The primary aim of the studies was to evaluate the effects of RP on school and classroom climate in middle schools, with consideration given to the potential moderating role of students' victimisation experiences and their emotional and social inclusion. The latter two variables were included, as they have previously been associated with school and classroom climate (Wang and Degol, 2016). The same hypotheses and research questions applied to both studies:

- 1) Training teachers in RP has a positive effect on school and classroom climate, as assessed by the students' perspective of
  - (a) classroom climate as a superordinate construct of the relationships in class, (b) especially the relationships amongst students in class, (c) especially the relationships between students and teachers in class, (d) rigour-control in school, and (e) warmth in school.
- 2) Potential effects of the teacher training in RP on classroom climate are moderated by (a) students' victimisation experiences and/or (b) students' perceptions of inclusion.
  - a) As addressing victimisation experiences is complex (Olweus, 1997), we assumed that for students with severe victimisation experiences before the RP intervention no significant changes in the school and classroom climate can be seen within our measurement period as more time is expected to be needed here for RP to take effect.
  - b) We further hypothesized that greater emotional and social inclusion before the RP intervention will be associated with more positive school and classroom climate after the intervention.
- 3) Feasibility:
  - a) It is possible to recruit students in inclusive classes for a study evaluating the effects of a teacher training intervention on school and classroom climate.
  - b) The chosen measures are sensitive enough to capture potential treatment effects on school and classroom climate on such a short time-scale of a staggered implementation of RP training components.

## STUDY 1

### Methods

#### Design

Study 1 used a single group pre-post design with the objective to evaluate the feasibility of the recruitment procedures, the suitability of the questionnaires, and initial findings regarding our research questions, using a small sample. Pre-tests were conducted between March and May 2019, with post-tests taking place in July 2019.

#### Procedure

##### *School Recruitment for the Training in Restorative Practices*

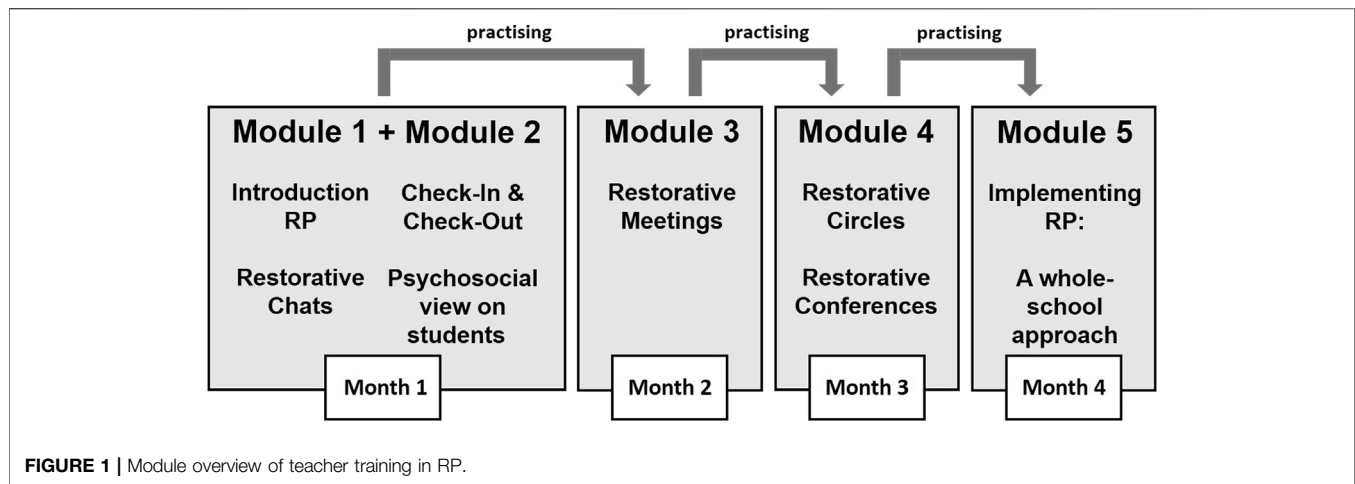
Inclusive secondary schools of a metropolitan area in the northwest of Germany were informed by email and telephone about the RP training offer. An information session for interested teachers and school administrators was held 2 months prior to the start of the intervention to clarify the objectives of the RP training and the research project.

The German secondary school types eligible to participate in the study were Hauptschule, Realschule, Gymnasium and Gesamtschule. A Hauptschule provides education for grades five through nine or ten. Likewise, the Realschule offers grades five to ten but is more practically orientated. A Gymnasium provides academic-oriented education to grades five to twelve or thirteen. Finally, the Gesamtschule encompasses all school previous types for grades five through twelve/thirteen.

##### *Participant Recruitment and Data Collection*

Teachers and educational staff from seven schools participated in the training. We contacted the teachers of the three schools reporting the highest proportion of children with SEND in their class and asked them to survey their pupils, all of the teachers agreed. On the first day of the RP training, we contacted these teachers to ascertain that they teach in grades 5 to 10. Subsequent to obtaining the consent of these teachers and their school management, students and their parents received study information sheets and written declarations of informed consent with the request to return these within 3 weeks, during which the researchers were available for questions or more information.

The questionnaires were administered during school hours and on the school premises by one or two researchers, depending on the number of participating students. Answering the questionnaires took about 45 min. Before answering any questionnaire the researcher stressed that participation is voluntary and further clarified the objectives and procedures

**TABLE 1 |** Demographic information of participants in Study 1.

Classes	<i>n</i> /class size	Age <i>M</i> ( <i>SD</i> )	Female %	SEND <sup>a</sup> ( <i>n</i> )	<i>n</i> at post-test
school 1	58/81	11.14 (1.01)	48.3	7	0
grade 5	22/27	10.45 (0.60)	45.5	3	0
grade 5	18/27	10.71 (0.47)	55.6	2	0
grade 7	18/27	12.39 (0.50)	44.4	2	0
school 2	38/67	12.63 (1.55)	47.4	11	37
grade 5	13/19	11.23 (0.44)	53.8	4	13
grade 6	12/22	12.17 (0.72)	41.7	5	12
grade 8	13/26	14.46 (0.88)	46.2	2	12
school 3	34/46	12.79 (2.07)	47.1	13	0
grade 5–7	20/24	11.40 (1.14)	50.0	5	0
grade 8–10	14/22	14.79 (1.25)	45.0	8	0
<b>Overall</b>	<b>8 classes</b>	<b>130/194</b>	<b>47.7</b>	<b>31</b>	<b>37</b>

<sup>a</sup>Of which 6 participants with a suspected SEND. The values in bold represent the values of the total sample across all three schools and all 8 grades.

for pseudo-anonymisation based on unique self-generated participant codes.

The scaling of the questionnaire items was illustrated using water glasses with different fill levels to make the scaling more comprehensible.

### Intervention: Training Teachers in Restorative Practices

Prior to developing the training manual of the current study, the research team received training in RP from internationally renowned trainers and institutes. The resulting training comprised five full-day training days spread over a 4-month period. **Figure 1** illustrates the training outline for the different modules (one per training day) and the practice phases between the training days. Further information regarding training content can be found in the supplementary materials (**Supplementary File S2**). Before taking part in the training, the participating teachers received approval from their school management to implement the training content.

### Participants

We recruited 130 students from three inclusive secondary schools (a Gymnasium, a Gesamtschule, and a Hauptschule) in a metropolitan area in the federal state of North Rhine-Westphalia in the northwest of Germany (**Table 1**). Classes were eligible to participate in the study if they included at least one student with SEND and at least one of their teachers, who had a minimum of five contact hours per week with the respective class, participated in the RP training. No exclusion criteria applied to the students of eligible classes.

### Instruments

*The Linzer Fragebogen Zum Schul- und Klassenklima (8–13)* Single scales of the Linzer Fragebogen zum Schul- und Klassenklima for the 8th to 13th grade (LFSK 8–13; Eder, 1998) were used to measure students' self-reported primary outcomes related to school and classroom climate. According to the manual, the questionnaire may also be applied in lower grades if there are instructions for answering the questions (Eder, 1998).

The scale capturing *relationship amongst students* contains 12 items and consists of two subscales: ‘community’ (six items capturing the degree of cohesion and mutual sympathy among students) and ‘rivalry’ (six items assessing the extent to which each student strives for achievement and success at the expense of their classmates). On the emotional level, rivalry means that one’s own success is valued higher when it is connected with the failure of others, or in the extreme case, that the failure of others represents a value in itself. Items on each subscale are scored on a Likert-scale of 1 (“not true”) to 5 (“exactly right”). Cronbach’s Alpha for the respective subscales is 0.74 and 0.80 (Eder, 1998) and therefore considered reliable (Bortz and Döring, 2006).

The scale *relationship between students and teachers* consists of five subscales, of which the following three subscales were used: ‘pedagogical engagement’, ‘restrictiveness’, and ‘injustice’. The subscales comprised six items each, with respective internal consistencies (Cronbach’s alpha) of 0.77, 0.78 and 0.78, (Eder, 1998), which was considered to be acceptable (Bortz and Döring, 2006).

We also used the single subscale “disruptions”, which measures the level of restlessness and disturbance caused by students in the classroom, with high levels considered indicative of a lack of concentration as well as disinterest in the classroom, and a low value indicative of a disciplined working atmosphere. The subscale consists of six items with a Cronbachs Alpha of 0.70 (Eder, 1998), which is considered acceptable (Bortz and Döring, 2006).

The scales *rigour-control* and *warmth* refer to the school as a whole. Rigour-control consists of six items, with an acceptable Cronbach’s Alpha of 0.76, whereas the warmth subscale consists of nine items with a solid Cronbach’s Alpha of 0.88 (Bortz and Döring, 2006).

For Study 1 and 2, the scores of each of these LFSK-scales were combined into an overall ‘*classroom climate scale*’.

### ***The Revised Peer Experience Questionnaire***

We measured experiences of victimisation among peers using nine items spread over three scales (three items each) of the victim version of the Revised Peer Experience Questionnaire (R-PEQ; De Los Reyes and Prinstein, 2004). This self-reported questionnaire assesses how often students experienced overt, relational and reputational aggression directed towards them within the past 3 months (e.g., “A teen chased me like he or she was really trying to hurt me”). Each item was coded to indicate how often (1 = never, 2 = once or twice, 3 = a few times, 4 = about once a week, 5 = a few times a week) each behaviour had been directed towards the informant. A sum score was calculated for each scale.

The original reliabilities of the scales are  $\alpha = 0.78$  for overt victimisation,  $\alpha = 0.84$  for relational victimisation and  $\alpha = 0.83$  for reputational victimisation (De Los Reyes and Prinstein, 2004), concluding that Cronbach’s Alpha is solid to acceptable (Bortz and Döring, 2006).

### ***Perceptions of Inclusion Questionnaire***

The emotional well-being in school and the social inclusion in class subscales of the Perceptions of Inclusion Questionnaire

(PIQ; Venetz et al., 2015) were used as additional indicators of school and classroom inclusion. Each scale comprised four self-report items to be rated from 0 = ‘not at all true’ to 3 = ‘certainly true’ (e.g., emotional inclusion: “I like going to school.”, social inclusion: “I have a lot of friends in my class.”) and sum scores were calculated for each scale. The reliabilities of the scales are  $\alpha = 0.90$  for emotional inclusion and  $\alpha = 0.83$  for social inclusion (Zurbriggen et al., 2017), and therefore considered excellent and solid (Bortz and Döring, 2006).

### ***Special Educational Needs and Disabilities Status***

Teachers provided written information regarding students’ SEND status in accordance with the current list of SEND recognised by the federal state of North Rhine-Westphalia (AO-SF NRW, 2016) and only when parents provided their informed consent for this action. To ascertain anonymous responding, students within a class were assigned a one-off number. Teachers used the one-off numbers to register which students had a particular SEND and which students they suspected to have a particular SEND. The one-off number allowed us to link SEND status to particular participants and even to non-participating students.

### **Plan for Analysis**

Initial baseline analysis included the calculation of intra-class correlations (ICC) to determine the proportion of the variance in the observations that lies between the classes. Given the distribution of the data, Mann-Whitney-U-tests were used to analyse differences in school and classroom climate, victimisation experiences, and social and emotional inclusion between pupils with and without SEND. Bonferroni-Holm corrections were applied to account for a potential  $\alpha$ -cumulation due to multiple comparisons.

Hypothesis 1 (training effects) was evaluated using Wilcoxon tests with Bonferroni-Holm corrections. As the test assumptions for an ANCOVA were not met and a random intercept model was considered inappropriate due to the small sample size ( $n = 1$  class at post-test; Maas and Hox, 2005; Schoppek, 2015), hypothesis 2 (moderator effects) was evaluated using non-parametric partial correlations to evaluate the relationship between school and classroom climate and experiences of victimisation, perceptions of inclusion, and the presence of SEND, whilst controlling for class membership. For further multiple linear regressions, we defined grade 5 of school 2 as the intervention group and grades 6 and 8 of this school, whose teachers dropped the training, as the control group. The multiple linear regressions were used to test whether the effect of teachers participating in RP intervention on the change in z-standardised classroom climate (dependent variable) is moderated by group membership and z-standardised victimisation experiences or the z-standardised social and emotional inclusion at pre-test (at group level).

All statistical analyses were performed using SPSS (International Business Machines Corporation (IBM), Armonk, NY, United States), Statistics for Windows, Version 27.0, considering  $p < 0.05$  to be significant.



**TABLE 2** | ICC, means and standard deviations of the outcome measures.

Scale	<i>n</i>	<i>M</i>	<i>SD</i> <sub>within</sub>	<i>SD</i> <sub>between</sub>	ICC
classroom climate	127	122.66	19.80	2.36	0.02
relationship amongst students	129	35.26	4.52	1.95	0.13
relationship between teachers and students	128	64.06	11.82	0	0
school climate					
rigour-control	128	20.81	3.33	2.06	0.28
warmth	126	33.68	6.13	0.59	0.01

Note. *SD*<sub>within</sub> = Standard deviation within the classes. *SD*<sub>between</sub> = Standard deviation between the classes.

## Ethical Approval and Consent

This study was approved by the Ethics Committee of Bielefeld University (EUB) (Approval ID: EUB 2019–005-A). Parents and students received study information and consent forms in different languages to improve the accessibility of the recruitment material.

## Results

### Participants

School 1 participants ( $n = 58$ ) were not surveyed at post-test as their teachers had not been able to implement methods from the RP training due to restructuring processes at the school-level. For participants from school 3 ( $n = 34$ ), we were informed during the study that class composition changes completely with each new school year and that therefore the class composition at pre-test would not be maintained at post-test. This would have made a direct pre-post comparison of the classroom climate—our primary outcome—impossible. Finally, for school 2, one student who participated in pre-test was absent due to illness on the day of the post-survey. Meanwhile, teachers of grade 6 ( $n = 12$ ) and 8 ( $n = 13$ ) from school 2 dropped out of the training after training day 2 without further information. Hence, reliable post-data were only collected for grade 5 students of school 2. This included 13 students, with a mean age of 11.23 years ( $SD = 0.44$ ), four of which had SEND (Table 1). Taken together, even if more pupils actually participated at post-test this represents a major pre-post attrition rate of 90%. When only taking school 2 into account, the attrition rate was still 66%.

### Baseline Data

Data from all eight participating classes ( $n = 130$ ) were included in the analysis of the baseline data. Taken all items together, ICC indicated 8.8% of the inter-individual variation in student perceptions was due to between class variability. Table 2 presents the ICC for each outcome measure. Due to the small sample, it was not possible to calculate a multi-level model to account for the class-level variability in this scale and data were analysed at class-level instead.

Mann-Whitney-U-tests with Bonferroni-Holm corrections yielded no significant differences between students with and without suspected or diagnosed SEND (across all classes) on either of the outcome measures (Supplementary File S3). The descriptive data suggest that students with SEND rate the outcome measures more negatively than children without SEND (Supplementary File S3).

### Hypothesis 1 (Pre-post Changes in School and Classroom Climate)

A non-parametric Wilcoxon test with Bonferroni-Holm-procedure was conducted for grade 5 of school 2 ( $n = 13$ ). The test yielded no significant differences between pre- and post-test-scores at class level (classroom climate:  $z = -1.827$ ,  $p = 0.068$ ; relationship amongst students:  $z = -0.707$ ,  $p = 0.480$ ; relationship between students and teachers:  $z = -1.736$ ,  $p = 0.083$ ; rigour-control:  $z = -1.250$ ,  $p = 0.211$ ; warmth:  $n = 12$ ,  $z = -2.803$ ,  $p = 0.005$  with a threshold of  $p < 0.001$  according to Bonferroni-Holm procedure), thereby rejecting hypothesis 1.

### Hypothesis 2 (Moderation Effects)

Partial correlations between school and classroom climate, experiences of victimisation, and perceptions of inclusion were calculated across all classes at pre-test, whilst controlling for class membership. As can be seen in Table 3, both the emotional and social inclusion scales of the PIQ and the different types of victimisation experiences from the R-PEQ were correlated with different LFSK-scales capturing student-teacher and student-student relationships, indicating that the scores on the different scales are related to each other.

Moderated multiple linear regressions were used to test whether the effect of teachers participating in RP intervention on the change in z-standardised classroom climate is moderated by z-standardised victimisation experiences or the z-standardised emotional and social inclusion at pre-test (at group level). For this purpose, we defined grade 5 of school 2 as the intervention group and grades 6 and 8 of this school, whose teachers dropped the training, as the control group. None of the five regression models reached significance (emotional inclusion:  $F(3,30) = 1.22$ ,  $p = 0.320$ , social inclusion:  $F(3,30) = 0.87$ ,  $p = 0.470$ , overt victimisation  $F(3,29) = 1.74$ ,  $p = 0.182$ , relational victimisation:  $F(3,29) = 2.71$ ,  $p = 0.063$ , reputational victimisation  $F(3,29) = 0.63$ ,  $p = 0.601$ , thereby rejecting hypothesis 2.

### Hypothesis 3 (Feasibility)

a) It is possible to recruit students in inclusive classes for a study evaluating the effects of a teacher training intervention on school and classroom climate.

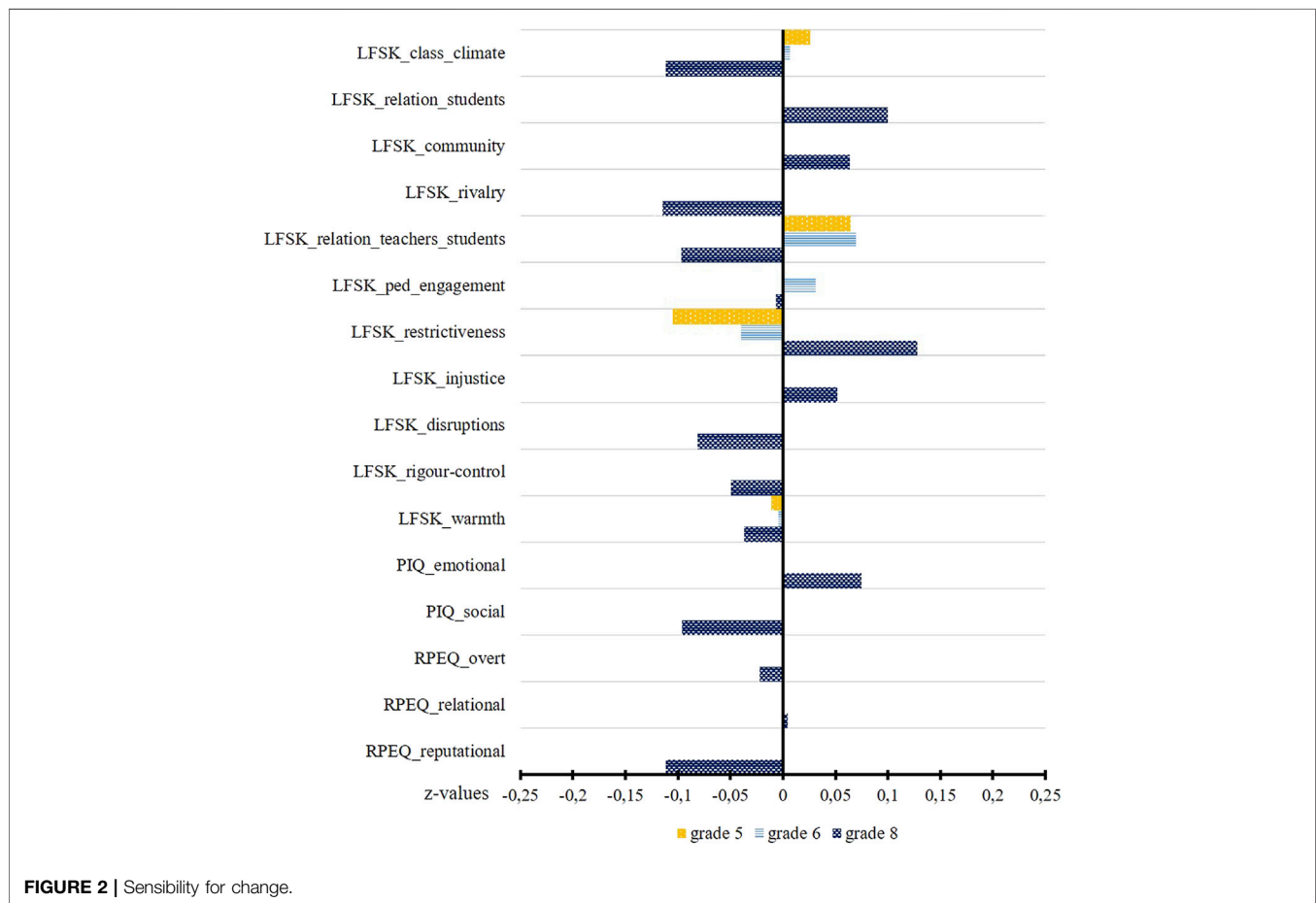
Across all eight participating classes the student participation rate was 67.01% at pre-test, ranging from 50 to 83.3% per class, with 23.85% ( $n = 31$ ) of participating students reportedly having

**TABLE 3 |** Spearman partial correlations between aspects of school and classroom climate and victimisation experiences, perceptions of inclusion and the presence of SEND.

LFSK	PIQ social inclusion	PIQ emotional inclusion	R-PEQ overt victimisation	R-PEQ relational victimisation	R-PEQ reputational victimisation	SEND <sup>a</sup>
<b>classroom climate</b>	0.27**	0.50**	-0.23*	-0.31**	-0.41**	-0.07
<b>relationship amongst students</b>	0.37**	0.22**	-0.12	-0.25**	-0.36**	-0.16
community	0.31**	0.18*	-0.04	-0.13	-0.18	-0.15
rivalry	-0.30**	-0.20*	0.18*	0.30**	0.45**	0.17
<b>relationship between students and teachers</b>	0.20*	0.57**	-0.20*	-0.27**	-0.38**	-0.04
pedagogical engagement	0.22**	0.54**	-0.04	-0.05	-0.10	-0.05
restrictiveness	-0.13	-0.41**	0.23**	0.24**	0.37**	0.05
injustice	-0.16	-0.45**	0.18*	0.32**	0.43**	0.09
disruptions	0.02	-0.23*	0.15	0.16	0.23*	-0.01
<b>school climate</b>						
<b>rigour-control</b>	0.08	0.11	-0.01	-0.22*	0.01	-0.17
<b>warmth</b>	0.22*	0.65**	-0.15	-0.19*	-0.31**	-0.04

Notes. \*indicates  $p < 0.05$  (two-tailed), \*\*indicates  $p < 0.01$  (two-tailed); control variable: class membership.

<sup>a</sup>Including the pupils with a suspected SEND.



SEND (including  $n = 6$  participants with suspected SEND), and 76.15% ( $n = 99$ ) of students without SEND. Only one out of three schools was still participating at post-test due to scheduling problems (school 1) or changes in class composition at the post-test (school 3), so the school retention rate was 33.3%. Within school 2 the retention rate was again 33.3%, because only one out of three classes was retained as teachers of two other

participating classes had dropped out of the training before any training gains could have been made. On an individual student level this corresponded to a participant retention rate of 90% with  $n = 13$  out of 130 participants participating in both pre- and post-test assessments. Within the participating class, 100% of the students participated again at post-test assessment.

- b) The chosen measures are sensitive enough to capture potential treatment effects of a staggered implementation of RP training components on school and classroom climate within a short time-frame.

Standardized mean differences for each class of school 2 were suggestive of pre-post differences, as illustrated in **Figure 2**, with negative values indicating a decrease and positive values representing an increase in the scores on the outcome measures. The chosen measures seem sensitive enough to capture fluctuations in school and classroom climate over time. Most changes, both positive and negative, have taken place in class 8, whose teacher stopped the training after training day 2.

## Discussion

Due to the small sample size, further reduced due to teacher drop-out in the training or school-based factors preventing post-test assessments, the analysis possibilities of Study 1 were limited. The available data did not provide evidence for the effects of teacher training in RP on classroom climate (hypothesis 1), nor for the presence of any moderator effects (hypothesis 2). The constructs studied correlated with each other to some extent, with student-student-relationships and teacher-student-relationships both being significantly related to students' perception of inclusion and their experiences of victimisation by other students.

Furthermore, we found no significant differences at pre-test between students with and without SEND regarding their perceptions of their school and classroom climate, victimisation experiences and social and emotional inclusion.

In summary, participation rates at baseline suggest that it is feasible to conduct teacher training and survey the students of these teachers in parallel (hypothesis 3), even if the response rates of Study 1 are lower than those reported in other studies (cf. Schwab, 2016). By contrast, with 24% of the consenting participants having SEND the proportion of children with SEND who participated in the study is comparatively high (cf. Schwab, 2016; Crede et al., 2019), thereby indicating that it is feasible to recruit in inclusive classes. Post-test data, however, did raise questions regarding the retention potential in our study. As there was hardly any dropout in the classes where assessment took place at both pre-test and post-test, it can be assumed that low attrition rates were more likely to be associated with school-level factors and less related with individual participant factors. Comprehensive anticipatory planning of post-testing appears necessary. Since it was possible to detect changes in the classroom climate, the measurement instruments seem to be appropriate to make changes visible.

## STUDY 2

### Methods

#### Design

A between groups non-randomised pre-post design was conducted with pre-tests taking place from September to December 2019 and post-tests taking place from February to

May 2020. Randomised allocation to the intervention and the control group was not possible as the study was part of a larger project where teachers' selection for RP training was based on interest shown by schools and their teachers. Hence, the control group was recruited separately.

### Procedure

#### *School Recruitment for the Training in Restorative Practices*

We used the same procedure to recruit schools as described in Study 1.

#### *Participant Recruitment and Data Collection*

Participant recruitment procedures for the intervention group were the same as those described in Study 1. Hence, recruitment was guided by the teachers participating in the RP training and took place in a metropolitan area in the northwest of Germany.

Teaching staff of 21 schools participated in the training, twelve of which were inclusive secondary schools. Following an information event for the training and the study, the participating class teachers of these 12 secondary schools were contacted by e-mail prior to the first training day to organise data collection in 5th to 7th grade classes with which they have a minimum of five weekly contact hours. As students in two of these schools received a concurrent intervention to reduce hostile attribution amongst students, their data are not included in this study. One school had already participated in Study 1, but had different teachers enrolled in RP training in Study 2 and thus different classes from this school were recruited for Study 2 compared to Study 1.

To recruit a separate control group, we contacted 63 inclusive schools in East Westphalia, four of which agreed to survey their students. Otherwise, recruitment procedures for the control group did not differ from the intervention group: Consenting teachers provided the study information sheets and consent forms to their students, whom were given up to 3 weeks to have their parents or legal guardian provide informed consent to participate in the study.

Prior to the training and/or recruitment we asked all classroom teachers in both groups which interventions or rituals they regularly use. The data in the intervention and control classes did not differ from each other. No school has worked with RP or, more specifically, relationship-based programs. Almost every class has worked with the "Klassenrat" (Blum and Blum, 2006). The Klassenrat is intended to promote democracy in the student body by giving students space to deliberate, discuss and decide on topics of their own choosing with assigned alternating roles (e.g., recorder, announcer, timekeeper). These are topics that concern the life together in the class, including conflicts. In contrast to RP, the focus is not on relationship work, but on living together in a community of responsibility, planning and coordinated action.

Pre-test assessments were scheduled to take place directly before the first day of RP training for the participating teachers and at a similar time interval for classes in the control group.

The data collection procedures followed those from Study 1 but lasted approximately 15 min longer, with a total participation



**TABLE 4** | Demographic information of participants in Study 2.

	Classes/schools	n/class size <sup>a</sup>	Age M (SD)	Female %	SEND <sup>b</sup>	n at post-test <sup>c</sup>
control group	13/4	223/329	10.91 (0.69)	44.4	47	180
intervention group	11/5	178/247	11.68 (1.17)	48.3	26	41
overall sample	24/9	401/576	11.22 (0.99)	46.1	73	221

<sup>a</sup>Class size of one class is missing in each group.

<sup>b</sup>Of which 24 participants with a suspected SEND. Calculated on the basis of the available data on SEND status as this information of some students is missing.

<sup>c</sup>Refers only to the participants who had already taken part in pre-test.

time of 60 min, because participants in Study 2 completed additional measures as part of the overall project. The data regarding these additional measures are not presented in this study.

Due to COVID-19 pandemic-related school closures, post-testing was moved forward from nine to 5 months after pre-test. These restrictions further caused the post-test assessment to be administered online, using Qualtrics software (Provo, UT, Copyright © 2020 Qualtrics), instead of using a paper and pencil format. To improve the accessibility of the online assessment, all items were provided with audio alternatives. Teachers received the URL for the online questionnaires with the request to share it with their students. With the online survey we had little control over the post data collection. However, not all students could be contacted by their teachers via email. Schools closed 6 weeks after the fifth and last training day for the participating teachers. Students participated in the online survey within the first 2 months of school closings.

### Intervention: Training in Restorative Practices

In Study 1, teachers reported that the theoretical content provided to them during the training contributed considerably to their use of restorative methods. This led us to include a new training section on attachment theory and how this affects students with and without SEND. We further invited an additional international expert to help us review our training. Differences in the training format and contents between Study 1 and Study 2 are detailed in the supplementary materials (**Supplementary File S2**). Most notably from an organisational perspective, the first two training days in Study 2 were not delivered on consecutive days but instead with an 8 week interval between them. Both trainings introduced teachers to the same RP methods, however, a key difference between them is the extent to which wider theoretical perspectives (e.g., attachment theory, inclusion and psychosocial development) were discussed.

### Participants

The same eligibility criteria applied as for Study 1 with the additional limitation that only students from grades 5 to 7 should participate. 221 students in the 24 participating schools provided data at baseline, of which 178 students from five different schools were in the intervention and 223 students from four different schools in the control group.

In the intervention group, a whole school with three participating classes had to be excluded, because the baseline survey could only be conducted in February 2020. In the control

group, four classes from one school had to be excluded because they did not include any students with SEND. Data of four classes of the control group and data of five classes of the intervention group are therefore included in the study. As the demographic data in **Table 4** indicate, 14.6% of students in the intervention group had (suspected) SEND, compared to 21.1% for the control group. Information on SEND status was missing for six classes in the intervention group and one class in the control group as the teachers were not willing to share the information despite parental consent being available. **Supplementary File S4** presents a detailed overview of the demographic data for each class.

### Instruments

Study 2 included the same measures as Study 1: LFSK 8–13 to assess school and classroom climate, R-PEQ to assess peer victimisation experiences, PIQ to assess students' perceptions of inclusion, and class-level proportion of SEND students. However, due to the constraints of the overall project this study was a part of, not all questionnaires were completed by all participants. In the control group, the PIQ was not administered and the R-PEQ was only completed by 6th grade participants. Additionally, we have developed questions to capture the implementation of the methods.

### Teacher Report of Restorative Practices

At post-test teachers were asked nine additional questions to evaluate the practical relevance of the training. These questions referred to the novelty of RP ("The philosophy of the restorative practices approach was new to me"), the significance of RP for their work ("The philosophy of the restorative practices approach was significant for me"), and to the frequency with which they applied seven individual RP methods, including Check-In, Check-Out, Restorative Chat, Restorative Meeting, Restorative Circle, Proactive Circle and Restorative Conference (e.g., "How many times have you run a Check-In?"). The first two questions are scored on a Likert-scale of 1 ("I totally disagree") to 6 ("I totally agree"). The questions about the frequency of use of the methods were captured on a Likert-scale of 1 ("never"), 2 ("1 to 5 times"), 3 ("6 to 10 times") and 4 ("more than 10 times").

### Plan for Analysis

Due to the different study design and sample sizes, the procedure differs in parts from Study 1. Again, initial baseline analysis included the calculation of ICC to

determine the proportion of the variance in the observations that lies between the classes. Given the distribution of the data, Mann-Whitney-U-tests were used to analyse differences in school and classroom climate between intervention and control group. The effect size  $r$  was calculated as  $Z$  statistic divided by square root of the sample size ( $r = z/\sqrt{n}$ ) and interpreted in line with Cohen's classification (1992). Multilevel-analyses could not be conducted, because the sample size at class level ( $n = 41$  students from eleven classes) was too small (Maas and Hox, 2005; Schoppek, 2015) in the intervention group. Hence, hypothesis 1 (training effects) was examined with mixed ANOVAs to determine whether the RP training (intervention vs control group) had a significant effect on school and classroom climate. Hypothesis 2 (moderator effects) was evaluated using multiple linear regression models predicting the z-standardised difference between post and pre-test assessment of classroom climate as a dependent variable and group membership, z-standardised victimisation experiences and the interaction of victimisation experiences and group membership as independent variables.

All statistical analyses were performed using SPSS (International Business Machines Corporation (IBM), Armonk, NY, United States), Statistics for Windows, Version 27.0. Unless otherwise specified, significance testing was based on  $\alpha = 0.05$ .

### Ethical Approval and Consent

This study was approved by the Ethics Committee of Bielefeld University (EUB) (Approval ID: EUB 2019-005-A). The same procedures applied as in Study 1. To incentivise students to participate in the online post-test assessment, classes of which at least 70% of students participated received a class gift of 60 euros.

## Results

### Participants

Participants of one class from school 4 in the control group ( $n = 9$ ) were not surveyed at post-test due to school closures because of the COVID 19 pandemic and the school's refusal to participate in the online survey as opposed to the paper and pencil questionnaires at pre-test. The twelve remaining classes in the control group could be surveyed before the school closures. Meanwhile, in the intervention group no class could be surveyed in school at post-assessment due to the school closures. Hence, reliable post-data were only collected for  $n = 41$  participants of eight classes from five schools online in the intervention group. Taken together, this represents a major pre-post attrition rate of 44.9% in Study 2. The attrition rate in the control group was 19.7% and in the intervention group 77% (see **Table 4**).

### Descriptive Data

The participants in the intervention group were older than in the control group ( $U = 9,683.00, p < 0.001$ ). Gender was well balanced across both study arms ( $\chi^2(2) = 2.15, p = 0.342$ ). Mann-Whitney-U-tests were assessed to proof differences between intervention

and control group on school and classroom climate, and found one significant difference between the groups on the relationship amongst students and teachers scale of the LFSK ( $U = 13,836.50, p < 0.001, r = 0.22$ ). Students in the intervention group rated this scale higher ( $M = 69.73, SD = 10.08$ ) than students in the control group ( $M = 65.27, SD = 10.72$ ). The effect size is small (Cohen, 1992). All outcome measures are presented in **Supplementary File S5**.

Using Mann-Whitney-U-tests with Bonferroni-Holm corrections, we did not find differences between students with and without suspected or diagnosed SEND (across all classes) on the outcome measures (across all classes) (**Supplementary File S6**).

Similar to Study 1, ICC were calculated to determine the proportion of the variance in the observations that lies between the classes. Data from all 24 participating classes ( $n = 401$ ) were included in the analysis of the baseline data. Taken all items together, ICC indicated only 15% of the inter-individual variation in student perceptions was due to between class variability. **Table 5** presents the ICC for each outcome measure.

### Hypothesis 1 (Pre-Post Changes in School and Classroom Climate)

Due to the high dropout in the intervention group, it was initially checked for a systematic dropout. For this purpose, the mean values of the outcome variables at pre-test in the intervention group were examined for differences between the subjects who dropped out and those who participated at post-test. Descriptive data indicates students participating at pre and post-test-assessment rated the outcomes at pre-test slightly higher than students, who only participated at pre-test. However, Mann-Whitney-U-tests revealed that participants at pre- and post-test did not differ significantly after applying Bonferroni-Holm correction. There were no differences between participants at pre- and post-test and dropped participants regarding age ( $U = 1841.50, p = 0.416$ ), gender ( $\chi^2(2) = 1.51, p = 0.470$ ), and the presence of suspected or diagnosed SEND ( $\chi^2(1) = 0.22, p = 0.641$ ). Therefore, a systematic dropout in the intervention group cannot be confirmed. The statistical analyses are presented in **Table 6**.

One teacher from each school in the intervention group (School 5-9) participated in the survey on the practical relevance of the training and frequency of use of the methods. The responses are presented in **Table 7** and indicate that the relevance of the training is predominantly high, whereas the novelty of it received more varied responses. The frequency data indicate that methods have been used to a small extent in each school with small differences between schools.

Mixed ANOVAs were performed to examine whether the RP training (intervention vs control group) had a significant interaction effect on school and classroom climate. There was homogeneity of covariances for the scales relationship amongst students ( $p = 0.125$ ), relationship between students and teachers ( $p = 0.339$ ) and rigour-control ( $p = 0.487$ ), but not for classroom climate ( $p = 0.009$ ) and warmth ( $p = 0.005$ ) as assessed by Box's

**TABLE 5 |** ICC, means and standard deviations of the outcome measures.

Scale (LFSK)	<i>n</i>	<i>M</i>	<i>SD</i> <sub>within</sub>	<i>SD</i> <sub>between</sub>	ICC
classroom climate	373	129.12	16.57	6.70	0.14
relationship amongst students	392	43.10	7.15	2.49	0.11
relationship between teachers and students	386	67.26	9.96	3.46	0.11
school climate					
rigour-control	397	20.70	3.16	1.62	0.21
warmth	383	34.42	5.40	2.54	0.18

Notes. *SD*<sub>within</sub> = Standard deviation within the classes. *SD*<sub>between</sub> = Standard deviation between the classes.

**TABLE 6 |** Differences according to Mann-Whitney-U-tests between pre-only-participants and pre- and post-participants regarding school and classroom climate at baseline.

Dependent variable (scale of the LFSK)	Group	<i>n</i>	<i>M (SD)</i>	<i>U</i>	<i>p</i>	
school climate	rigour-control	only pre	135	20.90 (3.82)	2,638.50	0.651
		pre and post	41	21.28 (3.27)		
	warmth	only pre	132	34.34 (6.37)	2028.00	0.044
		pre and post	39	36.68 (5.05)		
classroom climate	only pre	128	130.35 (17.10)	1933.00	0.089	
	pre and post	37	135.77 (16.90)			
	relationship amongst students	only pre	133	42.65 (7.80)	2,503.50	0.742
		pre and post	39	43.09 (7.73)		
	relationship between students and teachers	only pre	131	68.96 (10.04)	2,130.00	0.046
		pre and post	41	72.20 (9.95)		

Note. Using Holm's sequential correction no difference reached significance.

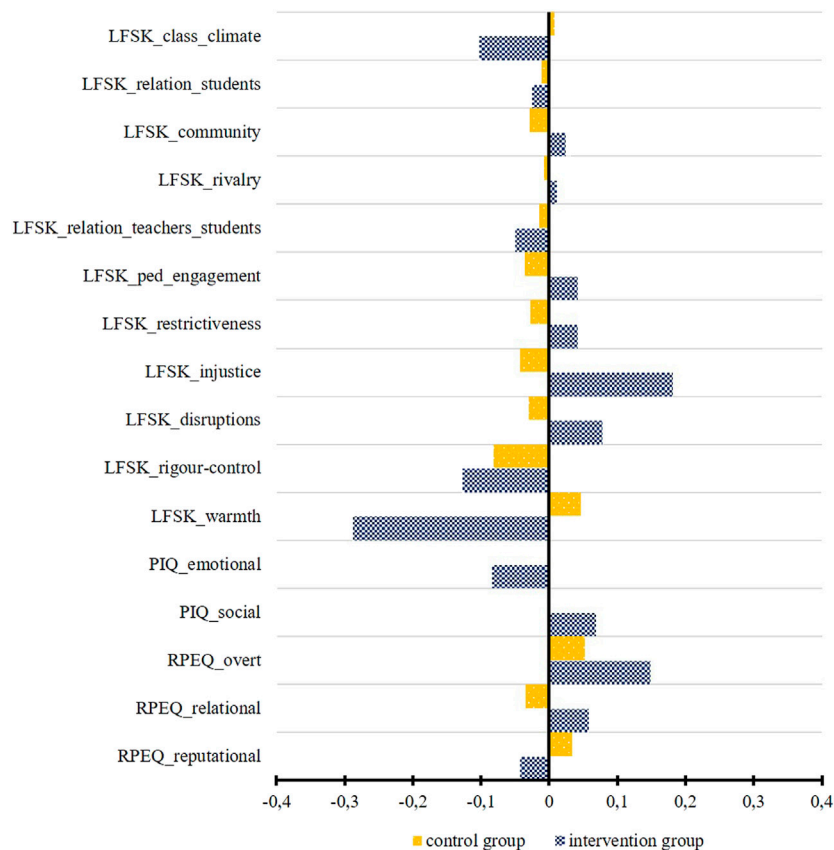
**TABLE 7 |** Survey of teachers about the practical relevance of the teacher training.

	School 5 teacher 1	School 6 teacher 1	School 7 teacher 1	School 8 teacher 1	School 9 teacher 1
novelty of RP (1 "totally disagree" to 6 "totally agree")	6	1	3	2	6
significance of RP (1 "totally disagree" to 6 "totally agree")	6	? <sup>a</sup>	5	4	6
Frequency of use of methods					
Check-In	0	1–5	6–10	1–5	1–5
Check-Out	0	? <sup>a</sup>	6–10	1–5	1–5
Restorative Chat	1–5	1–5	1–5	1–5	1–5
Restorative Meeting	0	? <sup>a</sup>	1–5	1–5	1–5
Restorative Circle	1–5	0	0	1–5	0
Proactive Circle	0	? <sup>a</sup>	6–10	1–5	1–5
Restorative Conference	0	? <sup>a</sup>	0	1–5	0

<sup>a</sup>Specification missing.

tests. Since the interaction effects between group membership and time of measurement are of central importance for answering hypothesis 1, we will only address these effects in the following. The statistical data on the main effects for group and time can be found in **Supplementary File S7**. Mixed ANOVAs regarding classroom climate ( $F(1, 189) = 2.38, p = 0.125$ ), the relationship amongst students ( $F(1, 213) = 5.53, p = 0.020$  with a Bonferroni-Holm correction considering  $p < 0.01$  to be significant), the relationship between students and

teachers ( $F(1, 215) = 0.68, p = 0.411$ ), rigour-control ( $F(1, 215) = 2.08, p = 0.151$ ), and warmth ( $F(1, 204) = 0.02, p = 0.885$ ) revealed there were no statistically significant interaction between time and group. Taken this results together, hypothesis 1 could not be confirmed. Nevertheless, the descriptive data of the assessment of most school and classroom climate aspects slightly deteriorated in the control group and slightly improved in the intervention group (**Supplementary File S8**).



**FIGURE 3 |** Sensibility for change.

### Hypothesis 2 (Moderation Effects)

Moderated multiple linear regressions were used to test whether group membership, z-standardised victimisation experiences at pre-test and the interaction of both predicted the z-standardised difference between pre- and post-test classroom climate ratings. The PIQ was not used in the control group, so a moderation effect regarding emotional and social inclusion on the pre-test cannot be verified. However, none of the regression models reached significance (overt victimisation:  $F(3,130) = 0.20, p = 0.894$ , relational victimisation:  $F(3,130) = 1.36, p = 0.260$ , reputational victimisation:  $F(3,129) = 0.31, p = 0.818$ , thereby rejecting hypothesis 2.

### Hypothesis 3 (Feasibility)

a) It is possible to recruit students in inclusive classes for a study evaluating the effects of a teacher training intervention on classroom climate.

Eight out of 12 schools in the intervention group (66.6%) agreed to the survey of their students. In the control group, 63 schools were contacted, four of which (6.3%) agreed to the survey. Information on SEND status was missing for six classes in the intervention group and one class in the control group. It is hence not possible to determine how many students

with SEND participated for these six classes. Across the remaining 22 classes the student participation rate was 69.62% at pre-test ranging from 21.74 to 100% per class. Overall, 18.2% ( $n = 73$ ) of the participating students reportedly had SEND (including  $n = 24$  students with suspected SEND), 63.6% ( $n = 255$ ) of students without SEND participated and 18.2% ( $n = 73$ ) participated with an unknown SEND status.

Both in the control and in the intervention group all schools retained until post-test. Post-test assessments were conducted shortly before the school closings due to the COVID-19 pandemic. One class could not be assessed again due to the school closures. Hence, within the schools of the control group the retention rate was 92.3%. This corresponded to a participant retention rate of 80.7% on an individual student level with  $n = 180$  out of 223 participants participating in both pre- and post-test-assessments. Within the participating classes, 91.9% of the students participated again at post-test assessment.

In the intervention group, post-test assessments could not be conducted before the school closures taking effect, leading to the shift to online post-test assessments. Within the schools of the intervention group the retention rate was 72.7% with eight out of eleven classes retained until post-assessment. This corresponded to a participant retention rate of 23% on an individual student level with  $n = 41$  out of 178 participants

participating in both pre- and post-test-assessments. In addition, eight students who participated at post-test could not be matched to any pre-test data and were not taken into account. Dropout for the intervention group accumulated to 77%. Within the participating classes, 25.9% of the students participated again at post-test assessment.

- b) The chosen measures are sensitive enough to capture potential treatment effects on school and classroom climate on such a short time-scale of a staggered implementation of RP training components.

Standardized mean differences for each group indicate pre-post differences, as illustrated in **Figure 3**, with negative values indicating a decrease and positive values representing an increase in the scores on the outcome measures. The assessment of school and classroom climate changed in the control group and in the intervention group between the survey times. In the control group the changes are considerably lower than in the intervention group. The chosen measures seem sensitive enough to capture fluctuations in school and classroom climate over time.

## Discussion

Due to the high dropout in the intervention group the analysis possibilities of Study 2 were limited since we could not implement multilevel-analyses. However, the available data did not provide evidence for the effects of teacher training in RP school and classroom climate (hypothesis 1). Turning to hypothesis 2, no moderator effects could be found, so hypothesis 2 could not be confirmed.

Participation rates at pre-test suggest that it is feasible to conduct teacher training and survey the students of these teachers in parallel, although it seemed rather difficult to obtain information on SEND status and it was rather difficult to recruit a control group. In the control group it was feasible to retain students to post-test-assessment, in the intervention group it was not. It appears the feasibility of the study was strongly influenced by the COVID-19 pandemic. The drop-out in the control group was much lower probably because the survey could still be realised in the classroom immediately before the school closings, whereas the post-survey in the intervention group had to be carried out online. Hence, hypothesis 3 can only be evaluated to a limited extent.

## GENERAL DISCUSSION

Contrary to our expectations, the implementation of a teacher training on RP did not lead to significant changes in the school and classroom climate from the students' perspective in either of the studies, thereby rejecting hypothesis 1. For Study 2, we did observe improvements in the school and classroom climate in the intervention group and found deteriorations in the control group. Meanwhile, hypothesis 2 was rejected as there were no moderator effects of victimisation experiences and/or perceived inclusion in either of the studies. Regarding the feasibility of the study designs, Study 1 and 2 indicated that recruiting schools and students in

combination with teacher training in RP proved feasible, but that problems arise with the retention of participants to post-test assessments. Participant attrition in both studies was strongly associated with school factors (restructuring, class composition changes) or external factors (COVID-19 pandemic-related school closures) rather than with individual students', teachers', or schools' motivation. Meanwhile, the choice of measures was considered to be suitable to capture potential changes in the classroom climate as changes could be observed in Study 2.

The findings of these and previous studies on RP indicate that it is not yet clear when or how to measure the effects of RP on classroom climate. The choice of sufficiently sensitive outcome measures is essential to capture any changes in the given time frame of a study, especially for those without follow-up data. It may generally be more helpful to also employ outcome measures that can track changes in class climate and interactions that are directly related to RP: such as, empathy or frequency of shaming situations.

The results of our study are in line with those of two currently conducted RCTs, noting however, that these studies were conducted in very different school contexts. For example, Augustine and colleagues (2018) examined the effect of RP in United States-American schools and reported lower classroom climate ratings for students of their intervention group compared to their control group. Likewise, Acosta and colleagues (2019) did not find any improvements in student ratings of school connectedness or school climate following a 2 year implementation of RP. Instead, they discovered that these outcomes were associated with students' self-reported experiences of how often their teachers used specific RP methods. It was unexpected, however, that students in the control schools reported to have experienced more RP (related to the frequency of use of specific methods) than would have been expected and only a minority of students in the intervention schools experienced RP to a great extent. Thus, if it is not specifically assessed to what extent RP are actually applied, the correct interpretation of the results is considerably more difficult. Meanwhile, in an earlier quasi-experimental pre-post study, Wong et al. (2011) found that the sense of belonging and school harmony of grade 7 to 9 students in Hong Kong decreased when they received no restorative interventions or received only partial RP. By contrast, those students receiving RP as part of a whole-school approach reported a slight increase in these outcomes. Together, these studies show that the level and extent of RP implementation requires further investigation. However, as both the RCT's were published when the current studies were already underway, their findings could not be considered in the design of either Study 1 or Study 2. To what extent the school cultural context affects the implementation and the effects of RP is unclear, as there are too few high quality international studies.

Despite the present study not finding any statistically significant evidence to support the importance of a restorative school environment for students, there were still indications of improvements in the expected direction. For example, the intervention classes in Study 2 showed improvements of classroom climate in the expected direction, while the classes



in the control group showed deteriorations. The results of both studies are inconsistent, however, as we did not find indications of positive changes in Study 1. Although the study design and sample size of Study 1 were generally weaker than for Study 2, the inconsistent findings could also be an indication that there are class-specific characteristics that make the application of RP more or less successful. In this regard, further research into these possible class-specific factors, using multilevel modelling analyses, would be desirable.

Similar to past evaluations of RP, combining the training and implementation of RP with a research project was challenging, in particular regarding the teachers' ongoing participation in the multi-day training and the teacher-researcher communication to organise the assessments. The COVID-19 pandemic was not foreseeable and presented us with additional challenges, including switching to administering the post-test assessments online. It is evident, that these circumstances have affected the quality of our data.

On the outcome level, school and classroom climate changed both in the intervention and in the control group. Our studies therefore confirm findings that school climate perceptions can evolve (Wang and Dishion, 2012). Since school processes seem to be dynamic and change continually (Wang and Degol, 2016), measuring school climate at one point in time may not be sufficient to explain patterns of change and increasing the time points for assessing this outcome measure are advisable.

## Strengths and Limitations

In this paper, we presented the first data regarding the implementation of RP in both a German educational context and focusing specifically on its implementation in inclusive classes. To that extent, we used a wide battery of self-report questionnaires to assess the outcomes from the students' perspective instead of assessing teachers' perspective of their students' in-class relationships. Due to intercultural differences with previous studies, that were mainly conducted in Anglo-American countries, our findings cannot readily be compared to previous reports. This lack of comparative data also impedes the ability to examine the extent to which school cultural factors may have an influence. Furthermore, both studies come with limitations that may impede the strength of our findings. First, selection bias is likely to have affected the composition of both the intervention groups and the control group, as training was offered on a voluntary basis and only a small proportion of schools contacted for the control group was willing to participate. This may have led to a selection bias amongst the participating schools and teachers potentially being more open or willing to implement new interventions, but also resulting in student samples that may not be representative for general German inclusive school population. Likewise, schools and students may have been more eager to participate in the training and survey when they knew they would do well. Further research with randomised allocation would be of help to reduce possible bias.

A further limiting factor was that in addition to students requiring parental consent to participate, we also required the approval of headteachers and class teachers to survey their students. However, some headteachers or class teachers

decided against letting their students participate due to ethical concerns, in spite of a positive review of the study by the University's ethics committee, or lack of time resources. Overall, the proportion of students surveyed in different classes varied as not all parents provided informed consent. Thus a participation bias on behalf of parents' consent cannot be excluded and is likely as it was observed that students with SEND were less likely to obtain informed consent to participate in the study. This decision to not participate in the study was therefore solely the parents' decision as we did not exclude any student with SEND who had parental consent.

Unfortunately, the information on the SEND status of some individual students was missing as teachers were unwilling to share the information despite having parental consent forms. This raises questions about the nature of the concerns teachers have about providing this information, but also indicates that researchers may need to obtain such information from the parents directly. It therefore seems exceedingly important for researchers to consider the contextual that determine how students would be able to participate in a study and to gather the views from all involved parties (parents, teachers and students).

Moreover, our implementation period has probably been too short to make changes visible (Wadhwa, 2015). However, other studies with a longer, 2-year implementation of RP (e.g., Augustine et al., 2018; Acosta et al., 2019) did not find a significant effect on school climate either. Hence, it is unclear over what time-frame, if any, RP affects school and classroom climate. As both school and classroom climate are very broad constructs, the potential effects of RP may become more tangible on more narrowly defined constructs that are not simultaneously influenced by many factors other than RP. Path-models of teachers' implementation of RP methods, students' perception of these methods, and student's perceptions of constructive relationships, self- and other needs, their own and others' behaviour in conflict situations, and alternative options of action in conflict situations could map the influence of RP on such narrow constructs.

Evidently, the study design of each study also affects the strength of our findings. For Study 1, the quality of the research methods as a single group pre-post design is considered weak, with the study's quality further reduced by its final sample size. Meanwhile, the methodological quality of Study 2 is stronger with a sufficiently large pre-test sample size. However, restrictions coming in place during the COVID-19 pandemic considerably reduced its sample size to where the statistical analyses suffered as a result and class-specific differences could not be taken into account. Due to the pandemic, external factors may have started having an influence that could have affected individual students in different ways. There is however no way to control for this with the lack of comparison data of schools that were not affected by the COVID-measures at the same time. The pandemic is likely to have had a major impact on the entire school day and also caused changes to the intended time-frame of Study 2. In order to keep the dropout rate as low as possible

despite the difficult circumstances, we provided audio explanations for each item, visualisations of the answer scales, a phone number for queries and a financial incentive for high participation at class level. Unfortunately, these efforts do not seem to have been sufficient to motivate a high level of participation in the online survey. Although desirable, we were unable to generate a personal request and repeated reminders of study participation for the students due to the lack of contact data (Smith et al., 2019). We relied on the teachers to support us and forward the survey to their students. Furthermore, research results indicate that a small financial incentive for each participant could have been more effective than a financial incentive provided to the class with a high response rate (Smith et al., 2019). Together, these circumstances negatively affected the replicability of our study. The replicability is further limited by the nature of the training intervention spanning multiple days. Despite having a training manual, many training components relied on direct input from the participating teachers regarding their experiences with implementing the RP methods.

Moreover, it remains unclear to what extent the teachers implemented the contents of the training in RP. We did intend to collect this information for Study 2, but due to the sudden move towards an online assessment of teachers at the start of the lockdown and teachers' sudden responsibilities to deliver distance education, their response rate was very low and the data is only of limited informative value (Table 7). Combining the findings of the post-test survey questions with reports of teachers during the training, we know that they have applied the content to some extent. However, we do not have observational data to support this and it is uncertain to what extent teachers chose to continue implementing RP after the end of the training. If the approach was not sufficiently applied, no effects could be found. This could be a neuralgic point explaining why expected effects have failed to materialise (cf. Acosta et al., 2019). As part of another study from the same research project, the teachers participating in the training were also interviewed in parallel to the survey of the school and classroom climate from the students' perspective. These interviews are likely to provide information on the extent to which RP has been applied and whether changes in the way conflicts are dealt with have occurred as a result, as well as the extent to which teachers' self-perceived competence has changed. To avoid a biased presentation of that parallel study's findings here, we refrain from presenting only a few supporting quotes and await the full findings.

Meanwhile, not all students will have experienced the same extent of RP methods: for example in a class of 25 it could be all experienced proactive methods but very few have made personal experiences with restorative chats or meetings. Hence, class level data on the implementation of restorative methods may not be as helpful.

Furthermore, it cannot be ruled out that some schools already apply some methods in advance, but only know them under a different name. The few responses of the teachers and their feedback during the training showed us predominantly that

the philosophy of the approach was new to them. Even though most of the teachers were roughly familiar with the proactive methods, e.g. circle talks, they were not familiar with the reactive methods as a consequence of an already existing conflict. However, the fundamental attitude of RP in the implementation of the methods and not the methods themselves are essential (Hendry, 2009).

A further limitation results from the operationalisation of inclusive education and SEND, as the terms are inconsistently defined (Grosche, 2015). As stated, we followed the definition of inclusive education and SEND of the federal state of North Rhine-Westphalia (Lütje-Klose et al., 2017), although we know that this definition of inclusive schooling as differentiated from integrative schooling is controversial (Wocken, 2009). In inclusion research, the group of children with SEND is usually examined in comparison to a group of children without SEND - this is also the case in our study. In the higher understanding of inclusion, however, this group no longer exists. Inclusion would then function fully if the grouping of children with and without SEND (two-group theory), which is perceived as stigmatizing, were to be dispensed with (Grosche, 2015). Moreover, we included students with a suspected diagnosis of SEND in this study. This decision was grounded in an educational-practice perspective that teachers can identify students who could benefit from additional support measures and that diagnosis of SEND can be delayed in many students. Further, in both studies, the proportion of students with SEND was significantly higher than the average proportion of students taught in inclusive schools with SEND in secondary education in NRW (9.1%), further giving rise to the suspicion of selection bias (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2020). Due to the sample size of our studies, it was not possible to examine potential differential effects of training on students with or without SEND status or control for the different types of SEND. Since research findings suggest students with behaviour that is considered to be problematic, as is often the case with emotional-social developmental disorder, are more likely to be socially excluded in inclusive schools (Bosse et al., 2018), it would be worthwhile to test, whether these children particularly profit from RP.

## Implications and Future Directions

Our findings raise further research questions as well as questions about intervention possibilities in the practical school context. The time between the implementation of RP and outcome assessment in our studies was limited to a period of five to 6 months. With regards to the implementation of RP, this can be considered a rather short time span to expect intervention effects to be observed (Wadhwa, 2015). Hence, our studies may not have captured the full effect of RP implementation on students' perceptions on classroom climate. In this respect, studies with significantly longer time spans are needed. Cluster RCTs with a follow-up design would be desirable to achieve robust results. This would also be in line with the dynamic nature of school and classroom climate (Wang and Degol, 2016).

Examining school climate longitudinally in a cluster RCT would elucidate how school climate changes as a result of a new program or system implementation.

Further on, assessing the fidelity of implementation of RP is needed to be taken into account when interpreting the results, and preferably from the students' and teachers' point of view. Moreover, future research could benefit from additional observational measures of both implementation and outcomes. Capturing the implementation process at classroom level would help to get a deeper insight of whether at all and how RP are applied. Accordingly, direct supervision of teachers during implementation is recommended to increase the quality of the training. This is necessary, because consistency and predictability of RP implementation are likely to affect the intervention's effectiveness. Prior findings indicate that the implementation of RP varies widely across schools (McCluskey et al., 2008). Consequently, the power and comparability of findings across studies is largely hampered, also because there does not appear to be a common definition of which methods are essential to RP (Daly, 2002; Sellman et al., 2014).

In our study, we did not aim for a whole-school implementation, as only individual teams of teachers from different schools participated in the training. Our decision was based on the fact that the whole-school implementation of RP could sometimes have a deterrent effect on school headteachers. Some researchers assume to achieve the best results, it takes commitment from the whole school staff (Hendry, 2009). With Roger's (2003) diffusion model of innovation in mind, to which also Thorsborne and Blood (2013) refer for a successful implementation of RP, we aimed our training at getting small teams of teachers excited about change and getting early adopters before moving to a whole-school approach. If we cannot even get early adopters, it would be difficult, if not impossible, to move to a whole-school approach. Due to the lack of effects, it would be important to investigate if a whole-school approach can have an effect on school and classroom climate as perceived by students.

## Conclusion

This study is the first controlled trial of the effects of a teacher training in RP on students' perceptions of classroom climate in inclusive secondary schools in Germany, which was able to generate first important insights. The studied 4-month intervention did not yield significant changes in the intervention group. The results were not significantly different from those of the control classes, but there was some descriptive evidence of deterioration in classroom climate in the control group and improvements in the intervention group. It was shown that it is feasible to conduct a teacher training in RP in Germany and to capture its impact at the student level. Since school and classroom climate are latent constructs that can be influenced by numerous other factors, cluster RCTs in inclusive schools in Germany with a follow-up design and narrow coverage of class-specific aspects are needed to further investigate the effects of training in RP on classroom climate.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Bielefeld University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

Conceptualising the studies: CW and LV. Data collection: CW. Data preparation: MR. Data analysis: CW, MR, and LV. Preparing and interpreting essential literature: CW and MR. Research ethics: LV and CW. Supervision of research processes to adhere to good scientific practice: LV. Writing main manuscript: CW. All authors shared responsibility for drafting of the work and final approval of the version to be published.

## FUNDING

The present study was funded by a grant to Bettina Amrhein, Stefan Fries, and LV from the German Federal Ministry of Education and Research (support code: 01NV1738, support line: Qualification of educational professionals for inclusive education (Qualifizierung der pädagogischen Fachkräfte für inklusive Bildung)). We acknowledge support for the publication costs by the Open Access Publication Fund of Bielefeld University. The sole responsibility for the article's contents lies with the authors.

## ACKNOWLEDGMENTS

We are grateful to the participating schools, teachers and students. In addition, the following persons provided support for the project: Bettina Amrhein, Benjamin Badstieber, Beatrice Brandt, Stefan Fries, Wiebke Ludewig, Annalena Sander, Alina Simonsen, Nora Schmidt and Luise Schröder.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2021.719357/full#supplementary-material>

## REFERENCES

- Acosta, J., Chinman, M., Ebener, P., Malone, P. S., Phillips, A., and Wilks, A. (2019). Evaluation of a Whole-School Change Intervention: Findings from a Two-Year Cluster-Randomized Trial of the Restorative Practices Intervention. *J. Youth Adolesc.* 48 (5), 876–890. doi:10.1007/s10964-019-01013-2
- Anfara, V. A., Jr, Evans, K. R., and Lester, J. N. (2013). Restorative justice in Education: What We Know So Far. *Middle Sch. J.* 44 (5), 57–63. doi:10.1080/00940771.2013.11461873
- AO-SF NRW (2016). Verordnung über die sonderpädagogische Förderung, den Hausunterricht und die Schule für Kranke (Ausbildungsordnung sonderpädagogische Förderung – AO-SF). Vom 29. April 2005 zuletzt geändert durch Verordnung vom 01. Juli 2016. Retrieved from: <https://bass.schul-welt.de/pdf/6225.pdf?20210104150207> (Accessed January 4, 2020).
- Augustine, C. H., Engberg, J., Grimm, G. E., Lee, E., Wang, E. L., Christianson, K., et al. (2018). Can Restorative Practices Improve School Climate and Curb Suspensions? an Evaluation of the Impact of Restorative Practices in a Mid-sized Urban School District. Retrieved from: [https://www.rand.org/pubs/research\\_reports/RR2840.html](https://www.rand.org/pubs/research_reports/RR2840.html) (Accessed February 12, 2021).
- Avramidis, E. (2013). Self-concept, Social Position and Social Participation of Pupils with SEN in Mainstream Primary Schools. *Res. Pap. Edu.* 28 (4), 421–442. doi:10.1080/02671522.2012.673006
- Blum, E., and Blum, H.-J. (2006). *Der Klassenrat. Ziele, Vorteile, Organisation.* Mülheim an der Ruhr: Verlag an der Ruhr.
- Bortz, J., and Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler.* 4th ed. Heidelberg: Springer.
- Bosse, S., Jæuthe, J., Lambrecht, J., Bogda, K., Koch, H., and Spörer, N. (2018). Die Sicht von Kindern auf Inklusion in der Schule. Die Entwicklung eines Messinstruments zur Erhebung der Einstellung zum gemeinsamen Lernen im Grundschulalter. *Empirische Sonderpädagogik* 10 (4), 329–345. doi:10.25656/01:16778
- Cambra, C., and Silvestre, N. (2003). Students with Special Educational Needs in the Inclusive Classroom: Social Integration and Self-Concept. *Eur. J. Spec. Needs Edu.* 18 (2), 197–208. doi:10.1080/0885625032000078989
- Cohen, J. (1992). A Power Primer. *Psychol. Bull.* 112 (1), 155–159. doi:10.1037/0033-2909.112.1.15510.1037//0033-2909.112.1.155
- Cohen, J., McCabe, L., Michelli, N. M., and Pickeral, T. (2009). School Climate: Research, Policy, Practice, and Teacher Education. *Teach. Coll. Rec.* 111 (1), 180–213.
- Costello, B., Wachtel, J., and Wachtel, T. (2009). *The Restorative Practices Handbook for Teachers, Disciplinarians, and Administrators: Building a Culture of Community in Schools.* Bethlehem, PA: International Institute for Restorative Practices.
- Crede, J., Wirthwein, L., Steinmayr, R., and Bergold, S. (2019). Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf im Bereich emotionale und soziale Entwicklung und ihre Peers im inklusiven Unterricht. *Z. für Pädagogische Psychol.* 33 (3–4), 207–221. doi:10.1024/1010-0652/a000244
- Daly, K. (2002). Restorative justice. *Punishment Soc.* 4 (1), 55–79. doi:10.1177/1462474022228464
- De Los Reyes, A., and Prinstein, M. J. (2004). Applying Depression-Distortion Hypotheses to the Assessment of Peer Victimization in Adolescents. *J. Clin. Child. Adolesc. Psychol.* 33 (2), 325–335. doi:10.1207/s15374424jccp3302\_14
- Eder, F. (1998). *Linzer Fragebogen zum Schul- und Klassenklima für die 8.-13. Klasse (LFSK 8-13).* Göttingen: Hogrefe.
- Eder, F. (1996). *Schul- und Klassenklima. Ausprägung, Determinanten und Wirkungen des Klimas an weiterführenden Schulen.* Innsbruck: StudienVerlag.
- Eder, F. (2018). “Schul- und Klassenklima,” in *Handwörterbuch Pädagogische Psychologie.* Editors D. H. Rost, J. R. Sparfeldt, and S. R. Buch (Weinheim und Basel: Psychologie Verlags Union), 696–707.
- Eder, F. (2002). Unterrichtsklima und Unterrichtsqualität. *Unterrichtswissenschaft* 30 (3), 213–229. doi:10.25656/01:7686
- Flook, L., Repetti, R. L., and Ullman, J. B. (2005). Classroom Social Experiences as Predictors of Academic Performance. *Dev. Psychol.* 41 (2), 319–327. doi:10.1037/0012-1649.41.2.319
- Frostad, P., and Pijl, S. J. (2007). Does Being Friendly Help in Making Friends? the Relation between the Social Position and Social Skills of Pupils with Special Needs in Mainstream Education. *Eur. J. Spec. Needs Edu.* 22 (1), 15–30. doi:10.1080/08856250601082224
- 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
- Green, A. E., Willging, C. E., Zamarin, K., Dehaiman, L. M., and Ruiloba, P. (2019). Cultivating Healing by Implementing Restorative Practices for Youth: Protocol for a Cluster Randomized Trial. *Int. J. Educ. Res.* 93, 168–176. doi:10.1016/j.ijer.2018.11.005
- Grewe, N. (2003). “Aktive Gestaltung des Klassenklimas Eine Empirische Interventionsstudie.” (Münster: LIT-Verlag).
- Grewe, N. (2017). “Soziale Interaktion und Klassenklima,” in *Lehrer-Schüler-Interaktion.* Editor M. Schweer (Wiesbaden: Springer VS), 547–560. doi:10.1007/978-3-658-15083-9\_24
- Grosche, M. (2015). “Was Ist Inklusion?,” in *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen.* Editors P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant, and M. Prenzel (Wiesbaden: Springer VS), 17–39. doi:10.1007/978-3-658-06604-8\_1
- Heisig, K. (2018). Bundesländerunterschiede im Förderschulsystem. *ifo Dresden berichtet* 25 (05), 10–16.
- Hendry, R. (2009). *Building and Restoring Respectful Relationships in Schools: A Guide to Using Restorative Practice.* Abingdon: Routledge.
- Huber, C., and Wilbert, J. (2012). Soziale Ausgrenzung von Schülern mit sonderpädagogischem Förderbedarf und niedrigen Schulleistungen im gemeinsamen Unterricht. *Empirische Sonderpädagogik* 4 (2), 147–165. doi:10.25656/01:9296
- Klemm, K. (2015). Inklusion in Deutschland. Daten und Fakten. Gütersloh: Bertelsmann Stiftung. Retrieved from: [https://www.unesco.de/fileadmin/medien/Dokumente/Bildung/139-2015\\_BST\\_Studie\\_Klemm\\_Inklusion\\_2015.pdf](https://www.unesco.de/fileadmin/medien/Dokumente/Bildung/139-2015_BST_Studie_Klemm_Inklusion_2015.pdf) (Accessed April 24, 2020).
- 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. Inclusive Edu.* 13 (2), 117–140. doi:10.1080/13603110701284680
- Koster, M., Pijl, S. J., Nakken, H., and van Houten, E. (2010). Social Participation of Students with Special Needs in Regular Primary Education in the Netherlands. *Int. J. Disabil. Dev. Edu.* 57 (1), 59–75. doi:10.1080/10349120903537905
- Koth, C. W., Bradshaw, C. P., and Leaf, P. J. (2008). A Multilevel Study of Predictors of Student Perceptions of School Climate: The Effect of Classroom-Level Factors. *J. Educ. Psychol.* 100 (1), 96–104. doi:10.1037/0022-0663.100.1.96
- Lütje-Klose, B., Neumann, P., and Strees, B. (2017). *Schulische Inklusion in Nordrhein-Westfalen (NRW) – sieben Jahre nach Ratifizierung der UN-BRK.* Zeitschrift für Inklusion. Retrieved from: <https://www.inklusion-online.net/index.php/inklusion-online/article/view/416> (Accessed May 5, 2021).
- Maas, C. J. M., and Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology* 1, 86–92. doi:10.1027/1614-2241.1.3.86
- McCluskey, G., Lloyd, G., Kane, J., Riddell, S., Stead, J., and Weedon, E. (2008). Can Restorative Practices in Schools Make a Difference? *Educ. Rev.* 60 (4), 405–417. doi:10.1080/00131910802393456
- Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (2020). Statistische Daten und Kennziffern zur Inklusion 2019/20-Statistische Übersicht Nr. 410. Retrieved from: [https://www.schulministerium.nrw.de/system/files/media/document/file/Inklusion\\_2019.pdf](https://www.schulministerium.nrw.de/system/files/media/document/file/Inklusion_2019.pdf) (Accessed January 16, 2021).
- Mirsky, L. (2007). SaferSanerSchools: Transforming School Cultures with Restorative Practices. *Reclaiming Child. Youth* 16 (2), 5–12.
- Olweus, D. (1997). Bully/victim Problems in School: Facts and Intervention. *Eur. J. Psychol. Educ.* 12 (4), 495–510. doi:10.1007/bf03172807
- Pijl, S. J., and Frostad, P. (2010). Peer Acceptance and Self-concept of Students with Disabilities in Regular Education. *Eur. J. Spec. Needs Edu.* 25 (1), 93–105. doi:10.1080/08856250903450947
- Rogers, E. M. (2003). *Diffusion of Innovations.* New York: Free Press.



- Schoppek, W. (2015). Mehrebenenanalyse Oder Varianzanalyse? *Z. für Entwicklungspsychologie Pädagogische Psychol.* 47 (4), 199–209. doi:10.1026/0049-8637/a000136
- Schulte-Körne, G. (2016). Psychische Störungen bei Kindern und Jugendlichen im schulischen Umfeld. *Deutsches Ärzteblatt* 113 (11), 183–190. doi:10.3238/ärztebl.2016.0183
- Schwab, S. (2016). Erfassung von sozialer Partizipation - Übereinstimmung zwischen Selbst- und Fremdsicht von Schülern mit und ohne sonderpädagogischen Förderbedarf. *Z. für Pädagogische Psychol.* 30 (4), 227–236. doi:10.1024/1010-0652/a000187
- Sellman, E., Cremin, H., and McCluskey, G. (2014). *Restorative Approaches to Conflict in Schools*. London: Routledge.
- Smith, M. G., Witte, M., Rocha, S., and Basner, M. (2019). Effectiveness of Incentives and Follow-Up on Increasing Survey Response Rates and Participation in Field Studies. *BMC Med. Res. Methodol.* 19 (1), 230. doi:10.1186/s12874-019-0868-8
- Spörer, N., Maaz, K., Vock, M., Schröder-Lenzen, A., Luka, T., Bosse, S., et al. (2015). Lernen in der inklusiven Grundschule. Zusammenhänge zwischen fachlichen Kompetenzen, Sozialklima und Facetten des Selbstkonzepts. *Unterrichtswissenschaft* 43 (1), 22–35. doi:10.25656/01:15522
- Thapa, A., Cohen, J., Guffey, S., and Higgins-D'Alessandro, A. (2013). A Review of School Climate Research. *Rev. Educ. Res.* 83 (3), 357–385. doi:10.3102/0034654313483907
- Thorsborne, M., and Blood, P. (2013). *Implementing Restorative Practices in Schools: A Practical Guide to Transforming School Communities*. London and Philadelphia: Jessica Kingsley Publishers.
- United Nations (2006). United Nations Convention on the Rights of Persons with Disabilities. Retrieved from: <http://www.un.org/esa/socdev/enable/rights/convtexte.htm> (Accessed March 28, 2019).
- Venetz, M., Zurbriggen, C. L. A., Eckhart, M., Schwab, S., and Hessels, M. G. P. (2015). *The Perceptions of Inclusion Questionnaire (PIQ)*. German Version. Retrieved from: <http://piqinfo.ch/wp-content/uploads/2019/08/piq-english.pdf> (Accessed February 24, 2021).
- Wadhwa, A. (2015). *Restorative justice in Urban Schools: Disrupting the School-To-Prison Pipeline*. London: Routledge.
- Wang, M.-T., and Degol, J. L. (2016). School Climate: a Review of the Construct, Measurement, and Impact on Student Outcomes. *Educ. Psychol. Rev.* 28, 315–352. doi:10.1007/s10648-015-9319-1
- Wang, M. T., and Dishion, T. J. (2012). The Trajectories of Adolescents' Perceptions of School Climate, Deviant Peer Affiliation, and Behavioral Problems during the Middle School Years. *J. Res. Adolesc.* 22 (1), 40–53. doi:10.1111/j.1532-7795.2011.00763.x
- Way, N., Reddy, R., and Rhodes, J. (2007). Students' Perceptions of School Climate during the Middle School Years: Associations with Trajectories of Psychological and Behavioral Adjustment. *Am. J. Community Psychol.* 40 (3-4), 194–213. doi:10.1007/s10464-007-9143-y
- Weber, C., and Vereenooghe, L. (2020). Reducing Conflicts in School Environments Using Restorative Practices: A Systematic Review. *Int. J. Educ. Res. Open* 1, 100009. doi:10.1016/j.ijedro.2020.100009
- Wocken, H. (2009). "Inklusion & Integration. Ein Versuch, die Integration vor der Abwertung und die Inklusion vor Träumereien zu bewahren," in *Integration und Inklusion auf dem Wege ins Gemeinwesen*. Editors A.-D. Stein, I. Niediek, and S. Krach (Bad Heilbrunn: Klinkhardt), 204–234.
- Wong, D. S., Cheng, C. H., Ngan, R. M., and Ma, S. K. (2011). Program Effectiveness of a Restorative Whole-School Approach for Tackling School Bullying in Hong Kong. *Int. J. Offender Ther. Comp. Criminol* 55 (6), 846–862. doi:10.1177/0306624X10374638
- Zurbriggen, C. L. A., Venetz, M., Schwab, S., and Hessels, M. G. P. (2017). A Psychometric Analysis of the Student Version of the Perceptions of Inclusion Questionnaire (PIQ). *Eur. J. Psychol. Assess.* 35 (5), 641–649. doi:10.1027/1015-5759/a0004

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Weber, Rehder and Vereenooghe. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.