frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 30 June 2023
DOI 10.3389/feduc.2023.1208854

OPEN ACCESS

EDITED BY
Melissa Christine Davis,
Edith Cowan University, Australia

REVIEWED BY
Carolyn Gentle-Genitty,
Indiana University Bloomington, United States
Sarah Kiefer,
University of South Florida, United States

*CORRESPONDENCE
Jannik Nitz
✉ jannik.nitz@uni-koeln.de

RECEIVED 19 April 2023
ACCEPTED 19 June 2023
PUBLISHED 30 June 2023

CITATION
Nitz J, Hagen T, Krull J, Verbeck L, Eiben K, Hanisch C and Hennemann T (2023) Tiers 1 and 2 of a German MTSS: impact of a multiple baseline study on elementary school students with disruptive behavior.
Front. Educ. 8:1208854.
doi: 10.3389/feduc.2023.1208854

Tiers 1 and 2 of a German MTSS: impact of a multiple baseline study on elementary school students with disruptive behavior

Jannik Nitz*, Tobias Hagen, Johanna Krull, Leonie Verbeck, Katrin Eiben, Charlotte Hanisch and Thomas Hennemann

Department of Special Education and Rehabilitation, Faculty of Human Sciences, University of Cologne, Cologne, Germany

Introduction: The prevalence of disruptive behavior is increasing worldwide, affecting the learning process and classroom climate, teacher–student relationships, and peer interactions. Disruptive behavior in childhood can have detrimental effects on a child's long-term growth and may predict similar behavior in adolescence and adulthood. Thus, it is not only contemporary education and school research that are concerned with addressing the issues of teacher workload reduction and improvement of students' learning environment, but it is also a socially relevant issue. The Multi-tiered Systems of Supports (MTSS) effectively mitigates some of these challenges in the education system and is recognized as a successful framework.

Methods: Therefore, this single-case study examined the first two tiers of an MTSS developed in Germany (Multimo) to determine the impact of its implementation on 32 students with disruptive behavior. The first hypothesis propose that the implementation of Tier 1 can effectively decrease disruptive behavior among all students and the second one that assigning students to Tier 2 can provide an additional reduction in disruptive behavior. Data analysis included multilevel analyses with piecewise-multilevel models and the overlap index, the Nonoverlap of All Pairs, based on teachers' direct behavior rating. All analyses focused exclusively on the outcome variable of disruptive behavior, which was estimated based on teachers' daily behavior ratings.

Results: The results of the study showed a significant reduction of disruptive behavior in elementary school students at Tier 1. Disruptive behavior decreased at Tier 2 as well but did not show a significant change compared with Tier 1.

Discussion: The study concludes that in the context of an MTSS, the Good Behavior Game and Daily Behavior Report Card can be used together.

KEYWORDS

MTSS, single-case study, multi-level analyses, disruptive behavior, elementary school, SWPBS, good behavior game, daily behavior report card

1. Introduction

One of the main concerns of teachers in different types of schools is low-disruption teaching and the associated development of action knowledge around appropriate interventions to address problem behavior (Liaupsin and Scott, 2007; Ruiz-Olivares et al., 2010; Herman et al., 2017). Additionally, the expression of disruptive behavior in childhood can have negative effects on the child's long-term development and predict disruptive behavior in adolescence and

adulthood (Fergusson et al., 2005; Loeber et al., 2009; Bradshaw et al., 2010). Therefore, behavioral problems that significantly disrupt teaching are been topics of pedagogical practice and research for a long time (Reed et al., 1998).

Furthermore, disruptive behavior in childhood has negative effects on a society's overall system of functioning (Cohen, 1998; Boldrini et al., 2023; Tehrani et al., 2023). For this reason, the reduction is not only a concern for teachers and parents but also has political and economic dimensions (Rivenbark et al., 2018). In this context, teachers are also required to have increased knowledge of dealing with disruptive behavior of students.

Addressing disruptive behavior in schools and designing a systematic environment to reduce such behavior are key areas of school research (Adamson et al., 2019; Zuniga and Cividini-Motta, 2021; Serrano et al., 2023). Part of this field is the investigation of the interplay between school interventions and the effects of implementing systemic school concepts. It receives significant international attention and is backed by robust research results (Reed et al., 1998; O'connor and Hayes, 2020; Gagnon et al., 2022; Boldrini et al., 2023). However, a challenge lies in the fact that the research designs of many studies cannot accurately trace the exact developmental trajectories of individual students, which can be better accomplished through single-case studies (Wilbert et al., 2022). Tracking specific developmental trajectories offers advantages, particularly for students with disruptive behaviors, as it enables more assessment of the effectiveness and timing of interventions (Lipien et al., 2023). Therefore, it is important, from political, social, and societal perspectives, to explore the impact of different school interventions and their application on students in terms of disruptive behavior.

2. Students' disruptive behavior

Disruptive behavior is an umbrella term for different externalizing behaviors. It refers to actions or behaviors that interfere with the normal functioning of a group of students, and can assume multiple forms, including verbal outbursts, physical aggression, or refusal to follow rules or guidelines (Liaupsin and Scott, 2007). It is often viewed as a problem in educational settings, where it can impede the learning process for both the individual exhibiting the behavior and their classmates (Fossum et al., 2008; Loeber et al., 2009). Disruptive behaviors negatively impact the classroom climate (Reaves et al., 2018), teachers' social relationship with the class (Cui, 2022), and peer to peer interactions (Barth et al., 2004). Disruptive behavior in elementary schools poses significant challenges as it hinders learning (Lopez Jimenez et al., 2016; Schmerse and Zitzmann, 2021), creates conflicts among students (Yamasaki and Nishida, 2009; Krause et al., 2021), and disrupts relationships (Horn et al., 2021). It also results in missed instruction and difficulties for teachers in managing behavior and resources, impacting both educators and students. Addressing disruptive behavior early is crucial to prevent long-term social and academic difficulties (Whitten, 2009; Tehrani et al., 2023).

Approximately 5.7% of children and adolescents worldwide exhibit disruptive behavior (Polanczyk et al., 2015) and the estimated prevalence has been increasing (Barkmann and Schulte-Markwort, 2012; Klasen et al., 2017; Husky et al., 2018; Klipker et al., 2018; Ravens-Sieberer et al., 2022). Additionally, studies in many Western countries have reported significant emotional stress on teachers (Ozamiz-Etxebarria et al., 2021; Agyapong et al., 2022). The effects of disruptive behavior on teacher health can, therefore, be inferred based on a similar development of prevalence statistics and empirical school research (Kokkinos et al., 2005; Pines and Keinan, 2005; Herman et al., 2017).

3. Multi-tiered support for students with disruptive behavior

In this context, the leading approach to address disruptive behaviors is the Multi-tiered System of Supports (MTSS) (Reynolds, 1962; Simonsen et al., 2021; Agyapong et al., 2022). A key feature of this framework is the use of multiple interventions to effectively address and reduce the occurrence of the targeted behavior (Batsche, 2014). The explicit design of MTSS concepts varies internationally by school and education system (Steed et al., 2014; Nelen et al., 2020). Tiered prevention was first developed in the 1950s to address public health concerns (Institute of Medicine Committee on Prevention of Mental Disorders, 1994). Presently, they are widely adopted because they do not provide merely a one-dimensional answer and seek to link different areas of school support (Gage et al., 2018; Charlton et al., 2021). They are essentially characterized by an attempt to align a student's problem areas with the interventions of the approach. The MTSS, depending on their origin and design, can target student behaviors or academic achievement, that is, academic learning. Both may be supported depending on the school's format or resources (Fox et al., 2021).

The MTSS usually comprises three tiers (Stoiber and Gettinger, 2016): Initially, it reaches all students at the first tier. Support is implemented here for the benefit of all students. If students continue to exhibit problem behavior despite universal support, they are moved to the second tier. This tier is characterized by the fact that approximately 10–15% of the students in a class receive another appropriate intervention as support in achieving the target behavior (Huber, 2013). For this, the target behavior must first be determined, which is accompanied by a diagnostic process. If an identified at-risk student does not exhibit any change in behavior, he or she is moved to the third tier. Here, another 1–5% of the students receive an intervention to further reduce the problem behavior. These interventions are often systemic and can also include external support networks (Eagle et al., 2015).

However, all applied interventions of the MTSS should be evidence-based in terms of theory (Stoiber and Gettinger, 2016; Simonsen et al., 2021). In addition, the diagnostic process and teamwork are important components of the MTSS. Many studies and systematic reviews have demonstrated the effectiveness of the MTSS on different target variables, such as self-efficacy (e.g., Skinner et al., 2014), stress experience, and behavioral (e.g., Majeika et al., 2020) and learning-related effects (e.g., Gage et al., 2015) on students. Additionally, positive effects on variables other than disruptive behavior (O'connor and Hayes, 2020), such as school or classroom climate (Charlton et al., 2021), and internalizing behavior problems (Weist et al., 2018) can be derived from different studies.

4. Good behavior game and daily behavior report card as MTSS interventions related to disruptive behavior

The MTSS with a focus on behavioral issues (MTSS-B) would typically include interventions such as support for positive behavior, social skills training, and individualized behavior plans (Briesch et al., 2019). These interventions may emphasize reducing disruptive behavior and promoting positive behavior in the classroom.

One of these potential tier-one interventions is the Good Behavior Game (GBG) (Barrish et al., 1969). It is a widely used and well-researched classroom-wide intervention that employs positive reinforcement to promote positive behavior and reduce disruptive behavior (Joslyn et al., 2019). The standard arrangement of the GBG involves dividing a class into teams, setting rules, and implementing an interdependent group contingency to encourage students to follow the rules. Points (also called fouls) are given to teams each time a member breaks a rule, and rewards are given to teams with the lowest points (fouls). This approach, along with variations, has been extensively studied. The GBG reduces disruptive behavior and inattention in children with psychiatric disorders, as demonstrated in several single-case studies by Groves and Austin (2017), Joslyn and Vollmer (2020), and Donaldson et al. (2017). Another single-case study by Pennington and McComas (2017) demonstrated improvement in on-task behavior in children with attention problems. Groves and Austin (2019) also reported reduction in negative peer interactions in students with psychiatric disorders during play.

The positive effect on students (with behavioral problems) has been recently demonstrated in the form of single-case studies in Germany (Leidig et al., 2022; Hagen et al., 2023). Leidig et al. (2022) show significant higher effects of the GBG for students exhibiting disruptive problem behavior. In addition, Hagen et al. (2023) indicate that the GBG mitigates students' disruptive behavior over the school day and not only after playing the GBG in class. Both studies were conducted in German inclusive primary schools and involved approximately 20–30 students. However, research on the GBG with students who have emotional and behavioral disorder (EBD) diagnoses and histories of delinquency is still limited (Joslyn and Vollmer, 2020).

Second-tier interventions such as the Daily Behavior Report Card (DBRC) (Volpe et al., 2013) are implemented to improve positive behavior or reduce disruptive behavior (Briesch et al., 2019). The DBRC is a commonly used intervention to document and provide feedback on student behavior. It includes a clear target behavior, a scale for periodic judgment, a daily monitoring system, and communication between the student's teacher and home. The DBRCs are differentiated from other forms of behavior ratings in terms of their summative rating and emphasis on feedback. Studies have shown that the DBRCs are efficient and effective in improving academic skills and addressing behavioral issues (Vannest et al., 2010). To use evidence-based methods, educators seek efficient techniques to address student behaviors, and the DBRCs offer flexibility in implementation with minimal interruption to instructional time. A meta-analysis by Pyle and Fabiano (2017) comprising 40 single-case studies demonstrated the efficacy of the DBRC as an intervention for students with Attention Deficit Hyperactivity Disorder (ADHD). The study suggests that the use of higher-quality designs and special education classroom settings can lead to a more rapid and effective behavioral change. Based on these findings, school psychologists, special educators, and clinicians are highly recommended to incorporate the DBRCs in their approach to address on-task and disruptive behavior of children with ADHD.

Both interventions, the GBG and DBRC, are effective in reducing disruptive behavior and promoting positive behavior in the classroom. Their combination within the MTSS has been conceptualized for the German-speaking region but has not yet been evaluated (Hanisch et al., 2019).

5. Research questions and hypotheses

As shown disruptive behavior is an increasingly prevalent issue in the Western school environment. Although there are many educational strategies to address disruptive behavior, the MTSS has been found to be effective in reducing such behavior and involving multiple stakeholders to address school-wide issues. However, further research is required to determine the effectiveness of different interventions within the context of the MTSS. The GBG and DBRC have been identified as potential interventions that may be well-suited for use within the MTSS. The interaction of these different tiers within an MTSS in Germany appears to be novel in the context of research and innovative along the theoretical concept of MTSS in Germany (Hanisch et al., 2019). Often, studies remain rather superficial regarding the effectiveness of different interventions or the interaction of the different tiers within an MTSS. Furthermore, a review by Lee and Gage (2020) also demonstrated the effectiveness of tiered support in relation to student behavior. This leads to the assumption that it seems to make sense to integrate behavior modification measures into an MTSS and to investigate the outcome regarding disruptive behavior. Therefore, the first question is about testing the effectiveness of the GBG in reducing the disruptive behavior embedded in the MTSS at the first tier.

> *RQ1*: To what extent is the disruptive behavior reduced after the implementation of the Good Behavior Game at Tier 1?

Based on RQ1, testing Tier 2 compared to Tier 1 seems reasonable. Presently, there are few comparable studies using single-case studies to test the effects of Tiers 1 and 2 on disruptive behavior. This leads to the second research question.

> *RQ2*: What is the impact of the assignment to Tier 2 on the development of disruptive behavior compared with Tier 1?

Based on the formulated questions, the following hypotheses can be derived, which form the basis of this study for the evaluation.

Hypotheses (H):

> *H1*: The Good Behavior Game reduces the disruptive behavior of students exhibiting disruptive behavior.

> *H2*: Assigning students to Tier 2 can further reduce students' disruptive behavior.

6. The conceptual basis for this study: Multimo Tiers 1 and 2

The conceptual basis for this study is a German MTSS called *Multimo* (Hanisch et al., 2019), which was implemented at several elementary schools in the state of North Rhine-Westphalia in 2021 and 2022. It is an approach that makes teachers, parents, students, and school social workers part of the support, and thus, works in a multimodal manner. At the first tier, the GBG (e.g., Barrish et al., 1969) was implemented, and at the second tier, the DBRC (e.g., Volpe et al., 2013) was implemented in addition to the GBG. The third tier intervention is the SCEP Coaching (Schulbasiertes Coaching bei Kindern mit expansivem Problemverhalten, Hanisch et al., 2018), which is not part of this evaluation. The implementation was scientifically accompanied and evaluated based on a single-case study.

6.1. Tier 1

The first tier comprises the implementation of the GBG, which reaches all students in the class at the universal level. As the Tier 1 intervention, the GBG is a classroom-based intervention that uses mutual dependence of students to promote appropriate behavior in the class (Barrish et al., 1969). This well-researched intervention is considered effective in reducing disruptive behavior and building positive school experiences. Flower et al. (2014) identified 22 peer-reviewed articles evaluating the effectiveness of the GBG in reducing disruptive student behavior and enhancing learning. Results from two longitudinal studies showed the GBG's positive impact on health-related dimensions in primary school-aged boys with aggressive behavior. The average intervention effect was $d = 0.50$, with a significant and immediate reduction in challenging behavior (-20.38%, $p < 0.01$) after the GBG's introduction. Additionally, it improved social integration and peer acceptance/rejection. A meta-analysis by Chaffee et al. (2017) confirmed its effectiveness in supporting student behavior in regular schools. In this study, it was implemented for all students according to the concept at the first tier.

Students in a pre-defined small group receive rewards when they achieve previously defined behavioral goals together. Key features of the intervention include: (1) selection of goals and rules, identification of rule violations, (2) definition of rewards, (3) division of students into two or more teams during a time-defined game phase, (4) use of prompts for rule violations during the game phase, and (5) daily and weekly reward for teams with the fewest rule violations (Joslyn et al., 2014; Lastrapes, 2014). It, therefore, uses cognitive-behavioral strategies to reduce disruptive behaviors in class and positively reinforce appropriate behaviors. The team aspect also supports mutual support among students in adhering to rules.

The teachers were prepared for the implementation of the GBG in the form of full-day training by the moderators. All open questions regarding the implementation and application were satisfactorily addressed. The university and the facilitators were available for weekly reflection meetings during the implementation, and these were used as required.

6.2. Tier 2

In the second tier, the teachers also received training from the moderators and social workers. For the teachers, the DBRC was elaborated so that they could subsequently use it for the students who did not sufficiently benefit from the GBG in Tier 1.

Thereafter, the DBRC was implemented according to its original design, which has been proven to be an effective intervention for students in the school setting. The DBRC is a feedback form that uses a rating scale and point system to assess the student's behavior according to specific goals (Taylor and Hill, 2017). Its key components include the identification of target behaviors, daily assessment of their occurrence, and regular report-sharing with relevant individuals, such as parents, teachers, and students (Vannest et al., 2010). A teacher acts as the implementer and sets a point goal with the student, allowing them to choose a reward if the goal is achieved within a predetermined period (e.g., 30 min, a class period, a day, or a week). At the end of this period, the teacher and student review the DBRC, and the latter receives performance feedback and reinforcement if the goal is met. The DBRCs motivate desirable behavior through reinforcement and indirectly reduce disruptive behavior by not providing consequences for it (Volpe et al., 2013).

Teachers' behavioral goals and agreements with parents and the community were determined on an individual basis and fed back to the university team or facilitators as required. In this manner, the DBRC was implemented in different schools. Additionally, the necessary freedom in the design of the measures and their fit to the students and the environment could be achieved. The DBRC was administered to at-risk children in this sample throughout the period in addition to the GBG. Student behavioral goals were adjusted as required in this process.

7. Method

7.1. Participants and setting

The sample of this single-case study comprised 29 first- and second-grade students of 12 classes at 18 inclusive elementary schools in the federal state of North Rhine-Westphalia (Table 1). The different schools had three to five classes per grade and are located in a rural but urban-adjacent area. The students were selected by teachers based on their disruptive problems. The teachers were asked to identify the three most striking children in the class. The parents and students provided their consent to participate in the study and were informed about the study objectives as well as the interventions involved.

Based on teacher ratings, all 29 nominated students (100%) exhibited disruptive risks on either the total scale or the disruptive behavior scale. The risk assessment was based on the total score of the Integrated Teacher Rating (ITRF, Casale et al., 2018) with a cut-off score of ≥ 11, the *Oppositional/Disruptive* sub-scale with a cut-off score of ≥ 4, and the *Academic Productivity/Disorganization* sub-scale with a cut-off score of ≥ 9, following the manual outlined in more detail in Chapter 7.3.1. Most students (93%) in the sample were male (27) and only two students were female (7%). The average age of the students was 7.14 years (SD = 0.516, median = 7) and the age range was 6–8 years.

TABLE 1 Descriptive data for all 29 cases.

Case	Class	Gender	Age	$ITRF_{APP}$	$ITRF_{DOB}$	$ITRF_{Total}$
C1	a	2	7	13	4	17
C2	a	2	8	17	4	21
C3	b	2	7	8	8	18
C4	b	2	7	4	5	9
C5	c	2	8	6	5	11
C6	c	1	7	8	7	15
C7	c	2	7	8	6	14
C8	d	2	8	7	5	12
C9	e	2	7	12	1	13
C10	e	2	7	5	17	22
C11	f	2	6	13	17	30
C12	f	2	6	11	18	29
C13	f	2	7	19	16	35
C14	g	2	7	8	10	18
C15	g	2	7	2	5	6
C16	g	2	7	10	2	12
C17	h	2	7	5	7	12
C18	h	2	7	6	7	13
C19	h	2	7	3	10	13
C20	i	2	7	22	21	43
C21	i	2	7	5	8	13
C22	j	2	7	18	17	35
C23	j	2	7	6	5	11
C24	j	2	7	16	16	32
C25	k	2	8	5	6	11
C26	k	2	8	1	15	16
C27	k	1	8	14	14	28
C28	l	2	7	0	15	15
C29	l	2	7	10	2	12

ITRF APP, ITRF Academic Productivity/Disorganization sub-scale (cut off ≥ 9); ITRF DOB, Oppositional/Disruptive sub-scale (cut off ≥ 4); ITRG Total, ITRF Total Score (cut off ≥ 11).

7.2. Implementation

7.2.1. The multiplier concept

To implement the MTSS in elementary schools, a multiplier concept was established, which included three stages of professionalization. First, a group of 18 experienced (special education) teachers and school administrators were trained by university staff. These 18 so-called moderators subsequently trained the school staff regarding the implementation of the MTSS and accompanied this process throughout the entire survey period. In full-day training sessions (approximately five full days), the concept of tiered support, individual interventions, and the diagnostic methodology formed the content of the qualification at the schools. The process included weekly exchange meetings with the moderators and the university project team. The implementation of the interventions in the schools was continuously discussed. Similar train-the-trainer concepts have proven to be economical and effective for the dissemination of training content and for monitoring (Pearce et al., 2012). They have also been successfully used in school contexts in German-speaking countries (Behr et al., 2020).

7.2.2. Schedule

The time-frame for the implementation of the interventions and the overall concept was approximately 7 months. Tier 1 was implemented in early December 2021; Tier 2 in late January 2022; and Tier 3 in March 2022. Accordingly, the GBG was used for approximately the entire project period (7 months); the DBRC together with the GBG for approximately 5 months; and, in addition, the SCEP Coaching for the last 3 months. The study design impacted the commencement and duration of implementation. This implies that the schools did not commence the interventions simultaneously; the time-gap was as much as 3 weeks (Figure 1).

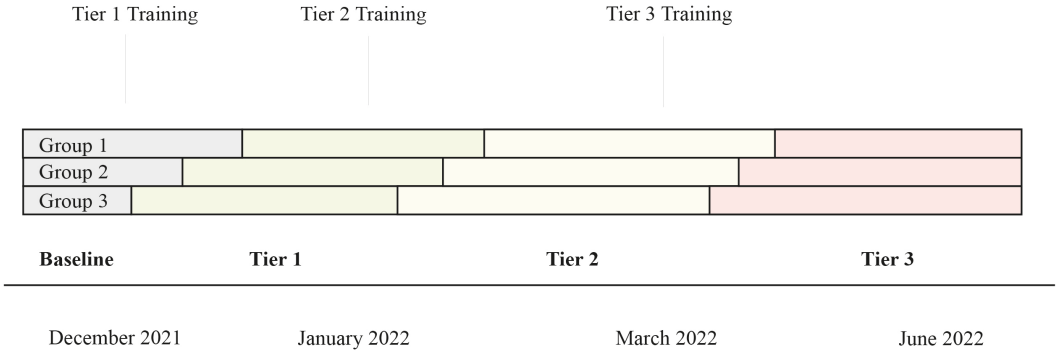

FIGURE 1
Schedule of the multiple baseline design.

7.3. Design and measures

7.3.1. ITRF

To screen children with disruptive behavior pre-selected by teachers, a short version of the German version of the Integrated Teacher Rating Form (ITRF) was used (Daniels et al., 2014; Casale et al., 2018; Volpe et al., 2020). The ITRF is a universal behavior screening tool that focuses on specific behavioral problems in educational settings. Its short version is divided into two sub-scales: *Oppositional/Disruptive,* and *Academic Productivity/Disorganization,* and includes 16 problem-oriented items that assess student behavior in the classroom, such as "Does not complete classwork on time," "Disrupts others," or "Moves around the room." The English ITRF meets common test validity criteria (Daniels et al., 2014); it has been translated into German and adapted to the German cultural context (Casale et al., 2018).

Casale et al. (2018) indicated high internal consistency for the oppositional and disruptive behavior (OD) scale ($\alpha = 0.96$) and the academic productivity and disorganization (APP) scale ($\alpha = 0.95$). Temporal stability over 2–4 weeks was acceptable with values of $r = 0.88$ (APP) and $r = 0.78$ (OD). The ITRF also demonstrated convergent validity with the Brief Problem Monitor-Teacher and divergent validity from the internalizing problem scales. Receiver operating characteristic curve analyses showed that the ITRF accurately identifies students at risk. This refers to students who, based on their exhibited behavior, should be assigned to either Tier 2 or Tier 3 within the decision-making process of an MTSS. Teachers have highly rated the ITRF for its perceived accuracy, feasibility, acceptability, and usefulness in informing interventions. In conclusion, the ITRF is a reliable and useful tool for teachers to identify and support students with behavioral issues in the classroom.

7.3.2. Direct behavior rating

Data were collected using the Direct Behavior Rating (DBR) (Huber and Rietz, 2015) with a scale on the extent of disruptive behavior (0–10). Disruptive behavior was defined as *Behavior that disrupts class or interferes with other children's learning. Examples: shouting in the class, fooling around, inappropriate side talk, or not staying in one's seat.* Along with this definition, teachers filled out the DBR. The definition was visible on the top of the DBR.

The DBR is a combination of systematic observation and rating and is deemed suitable for capturing detailed behavioral processes in the context of single-case studies because of its cost-effectiveness and versatility. It assessed behavior at a specific time during the school day. Huber and Rietz (2015) conducted a comprehensive analysis of its performance. They presented acceptable results regarding the reliability, validity, and accuracy of the instrument, although the results may vary depending on the type of behavior being evaluated and who is conducting the assessment. The study focused on the domain of learning and work behaviors, and the DBR produced reliable results, with an $\varepsilon 2$ and ϕ score of ≥ 0.80, after only four measurements (Volpe et al., 2012).

For this study, teachers were instructed to complete the DBR no later than 10 min after using the intervention (GBG). During the baseline survey, teachers completed the DBR once a day in comparable situations. The number of times the behavior was assessed varied from 83 to 102. In this manner, data were collected by operationalizing a concrete behavior and evaluating it in a specific scenario, always immediately after an explicitly defined time using rating scales. The disruptive behavior was rated using a Likert scale from zero to 10—zero representing low disruptive behavior and 10 indicating high disruptive behavior.

7.3.3. Single-case study

The single-case study was planned and conducted in a multiple baseline design along the presented schedule (Swoboda et al., 2010). For this purpose, the different schools were each given different starting times for the interventions. Consequently, the baseline phases differed from one another. The range of data points in the baseline was 15, 20, and 25. The data were entered conventionally using paper and pencil and the DBRs were prepared in advance by the project team and delivered to the teachers. To ensure that the teachers implemented the multiple baseline design along with the research design, DBRs contained explicit instructions that enabled the identification of the start times of the interventions visually and in writing. In addition, ongoing monitoring of the project team occurred. Thus, it was possible to ensure that the research design could be conveniently integrated into everyday school life.

7.3.4. Data analysis

Data analysis was conducted regarding interference statistics and descriptive statistics for multilevel analyses for single-case data. This includes piecewise-multilevel-models, measures, mean values, standard deviations, and necessary overlap indices. This form of data collection and analysis is recognized as a high-quality approach in the context of recent empirical interventions and school research. A distinctive feature of single-case studies is the students' own reference norm created by the baseline. In addition, the regression-based analyses meet the highest scientific standards, and therefore, offer a high degree of validity (Wilbert et al., 2022). The three phases comprised 1,498 measurement time points of all cases combined.

First, the cases were considered individually. For this purpose, the mean values in connection with the Nonoverlap of All Pairs (NAP) for Phases B and C were calculated and evaluated. In this process, the values were compared with their respective previous phases (A vs. B; B vs. C). The NAP (Parker and Vannest, 2009) is a commonly used index to estimate the overlap between different phases in single-case studies. To calculate the NAP, we first determine the number of data points that overlap between two phases in a single-case study. Next, we divide this value by the total number of pairs (data points) to obtain an index score between 0 and 1. Higher scores indicate greater nonoverlap, whereas lower scores represent more overlap. Additionally, the NAP considers any outliers or extreme values in the dataset that may affect its accuracy. A value in the range of 50–65 is considered a small effect; a value in the range 66–92 is considered a moderate effect; and a value over 93 is considered a large effect.

Additionally, piecewise-multilevel-regression analyses were calculated across all cases. In this manner, changes across the entire sample can be identified. Based on the piecewise-multilevel-regression analyses, level and slope effects were calculated. The level effect is an intervention effect that leads to an immediate and constant change in behavior at the beginning of the intervention phase. The slope effect is another intervention effect that continuously develops with the start of the intervention (Wilbert and Lueke, 2022). Both were set as fixed effects. R Studio (Posit-Team, 2022) was used for all analyses and here the scan package (Wilbert and Lueke, 2022) was used.

Owing to the complexity of the research questions and the associated significance, the analyses are limited to the first two tiers of the approach. Accordingly, two models were developed: the first facilitates statements about the GBG effect (Tier 1) as an intervention on the first tier, and the second allows statements about the influence of the assignment of a student to the second tier. For the first model, a multilevel regression was calculated across all cases, comparing the baseline phase with Phase B, that is, the intervention period of the GBG. For Model 2, the disruptive behavior of Phase B was compared with that of Phase C, which is the intervention period of Tier 2. No slope effect was calculated for Model 2 because of the focus on the phase difference from B to C. This was calculated in the form of the level effect from Phase B to C.

8. Results

8.1. Descriptive statistics

Table 2 summarizes the descriptive statistics for disruptive behavior, including the NAP as a measure of effect size (Alresheed et al., 2013). The non-rescaled NAP (comparing Phases A and B) values revealed that the dependent variable had effects ranging between 5.13 and 100%. More specifically, 13 participants (44.82%) showed small effects; 15 (51.72%) showed medium effects; and one showed strong effects.

The non-rescaled NAP (comparing Phases B and C) values revealed that the dependent variable had effects ranging between 27.4 and 99.8%. More specifically, 19 participants (65.51%) showed small effects; eight (27.58%) showed medium effects; and two showed strong effects (6.89%).

8.2. Inferential statistics

8.2.1. Model 1: comparing baseline and Tier 1 (RQ 1)

In Model 1, the disruptive behavior during three distinct phases is analyzed: Phase A (baseline with no intervention), Phase B (Tier 1), and Phase C (Tier 2), using a multilevel piecewise-regression model. The results show that the intercept of the regression line starts at 3.35. Additionally, the trend effect is significant and positive, with a value of 0.04. This effect includes all data points, even the baseline (Phase A), without any intervention. The level effect of both Phases B and C were significant when compared to the baseline (Phase A) (Table 3).

Specifically, the level effects of Phases B and C were −0.86 and −3.13, respectively. The level and slope effects of Phase C do not represent an adjusted effect, as they are influenced by Phase B. However, further investigation is required to verify the results of the level effect of Phase C. This is achieved by Model 2, which presents an adjusted value of the level effect of Phase C. The slope effects of both Phases B and C were significant. The slope effect of Phase B was −0.06, indicating that disruptive behavior is reduced by −0.06 per measurement time point. Similarly, the slope effect of Phase C was −0.08, indicating that disruptive behavior is reduced by −0.08 per measurement time point. Figure 2 is a schematic representation of Model 1 and its effects.

8.2.2. Model 2: comparing Tiers 1 and 2 (RQ 2)

In Model 2, we examined the disruptive behavior of Phases B and C. Our analysis was focused solely on these two phases. The trend effect was significant at −0.02, indicating a decrease in disruptive behavior in both phases together. However, our findings did not reveal a significant level effect between Phases B and C (Table 4). It only shows a decreasing effect in disruptive behavior, which is so small that the levels of the two phases do not significantly differ from each other. Across Phases B and C, disruptive behavior decreases along the significant trend effect (−0.02), but no significant level effect is evident to infer a significant change in behavior based on the mean comparison of both phases. Figure 3 is a schematic representation of Model 2 and its effects.

9. Discussion

The aim of RQ1 was to investigate the effectiveness of Tier 1 (GBG) in reducing disruptive behavior. The results show that through the implementation of the GBG, the disruptive behavior of elementary school students can be significantly reduced. The hypothesis (H1)

TABLE 2 Single-case data with NAP for all 29 cases.

Case	n_A	n_B	n_C	mis_A	mis_B	mis_C	M_A (SD)	M_B (SD)	M_C (SD)	NAP_{A-B}*	NAP_{B-C}*
C1	20	45	34	11	20	14	3.00 (0.50)	2.48 (1.08)	2.70 (0.98)	68.44	43.40
C2	20	45	34	16	20	17	1.75 (0.96)	1.60 (0.65)	0.47 (0.51)	53.50	88.71
C3	15	40	30	1	15	1	3.86 (2.88)	4.32 (2.23)	3.31 (1.44)	45.14	69.52
C4	15	40	30	1	18	2	2.71 (2.58)	2.00 (2.67)	0.29 (0.71)	60.71	69.48
C5	15	40	30	3	30	13	2.67 (1.97)	2.10 (2.18)	0.18 (0.53)	68.75	89.41
C6	15	40	30	12	24	12	4.67 (2.31)	1.62 (1.26)	1.33 (1.24)	89.58	56.42
C7	15	40	30	12	21	10	5.33 (0.58)	1.26 (1.05)	0.35 (0.49)	100	77.50
C8	15	40	43	6	32	10	6.00 (3.20)	2.38 (2.20)	1.52 (1.89)	81.94	65.91
C9	25	45	18	0	18	8	5.04 (1.67)	5.15 (3.36)	4.90 (2.96)	49.63	51.48
C10	25	45	18	3	21	15	6.32 (1.39)	4.75 (2.82)	3.00 (3.46)	65.81	73.61
C11	25	45	22	13	29	14	4.17 (1.90)	3.88 (2.06)	4.88 (1.81)	57.03	34.38
C12	25	45	22	10	29	13	4.87 (2.26)	1.69 (0.87)	2.11 (1.17)	90.21	39.58
C13	25	45	22	11	28	13	2.14 (0.95)	2.53 (1.46)	3.89 (2.52)	44.75	34.31
C14	25	45	30	12	18	5	1.77 (1.59)	6.08 (2.02)	2.28 (2.19)	5.13	88.83
C15	25	45	30	8	19	10	2.76 (3.07)	5.64 (0.92)	1.30 (1.13)	24.06	99.77
C16	25	45	30	8	18	15	3.53 (3.73)	4.50 (3.12)	0.00 (0.00)	42.65	91.67
C17	15	40	28	1	13	12	1.50 (0.85)	0.89 (0.70)	1.25 (1.00)	70.50	40.54
C18	15	40	28	1	13	13	1.43 (1.28)	0.44 (0.58)	0.00 (0.00)	75.66	70.37
C19	15	40	28	1	13	8	1.29 (1.20)	0.41 (0.75)	0.00 (0.00)	73.54	64.81
C20	25	45	32	3	20	3	9.86 (0.35)	7.44 (3.15)	8.72 (1.36)	73.64	45.31
C21	25	45	32	2	20	3	1.00 (0.52)	0.72 (0.68)	0.69 (0.71)	62.17	51.59
C22	25	45	25	5	26	10	7.40 (2.50)	5.32 (2.33)	7.13 (1.96)	75.26	27.37
C23	25	45	25	6	29	7	5.74 (3.90)	4.06 (2.93)	3.17 (2.43)	62.17	56.42
C24	25	45	25	7	24	11	6.28 (3.10)	6.05 (2.29)	7.36 (2.24)	55.69	33.16
C25	25	45	15	6	13	3	2.11 (1.73)	0.75 (0.92)	0.25 (0.45)	73.77	64.06
C26	25	45	15	6	12	7	5.00 (3.27)	2.79 (2.83)	3.50 (2.93)	69.30	40.91
C27	25	45	15	5	14	4	2.00 (2.53)	0.03 (0.18)	0.00 (0.00)	74.11	51.61
C28	25	45	15	3	23	11	6.50 (1.26)	5.05 (1.84)	6.00 (1.41)	73.35	34.66
C29	25	45	15	3	28	11	2.45 (1.65)	1.12 (1.54)	1.00 (0.82)	73.26	44.85

N, number of data points per specified phase; mis, missing data points per specified phase; M, mean of specified phase; SD, standard deviation; NAP, non-rescaled non-overlap of all pairs* is given in %.

TABLE 3 Fixed effects of the multilevel piecewise-regression model comparing the baseline with Phase B (Tier 1) and C (Tier 2) of the disruptive behavior.

Parameter	*B*	SE	*t*	*p*
Fixed effects				
Intercept	3.35	0.39	8.56	<0.001
Trend	0.04	0.01	3.48	<0.001
Level Phase B	−0.86	0.24	−3.58	<0.001
Level Phase C	−3.13	0.67	−4.67	<0.001
Slope Phase B	−0.06	0.01	−4.69	<0.001
Slope Phase C	−0.08	0.02	−4.93	<0.001

made in this context can, therefore, be confirmed. The decisive factors here are the key figures of the individual NAP comparisons between the baseline and Phase B, as well as the significant overall effect of −0.86 in the comparison of Phases A and B. This is consistent with the results of other studies conducted on German-speaking samples (Leidig et al., 2022; Hagen et al., 2023). These studies were also single-case studies, but here the GBG was not implemented in the context of the MTSS. Compared to the results of this study, the effects of these two studies were greater. This could be explained by the timing of the research: several schools had just gone through a lockdown period, which complicated various processes within the school. It is difficult to quantify the impact of this factor; however, it should not be overlooked.

In summary, there is an expected positive effect on the expression of disruptive behavior at the first tier of the MTSS across most cases. For more than half of the students (16), at least medium effects (NAP) were identified, and 13 students showed at least small effects. This, in addition to the data from the regression analyses, suggests an effect related to significant parts of the class.

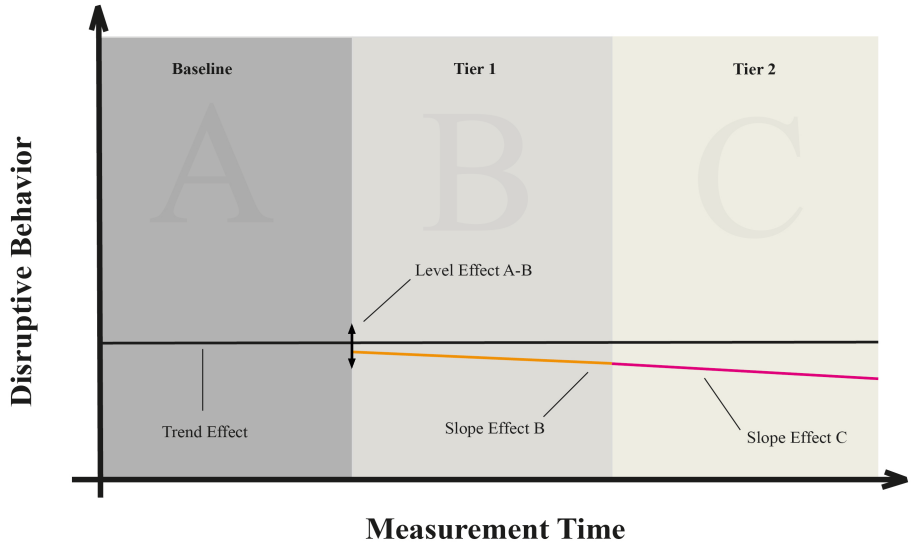

FIGURE 2
Schematic representation of Model 1.

TABLE 4 Fixed effects of the multilevel piecewise-regression model comparing Phase B (Tier 1) and C (Tier 2) of disruptive behavior.

Parameter	*B*	SE	*t*	*p*
Fixed effects				
Intercept	3.25	0.41	7.91	<0.001
Trend	−0.02	0.01	−3.16	<0.01
Level Phase C	0.03	0.21	0.16	0.87

The aim of RQ2 was to investigate the impact of assignment to Tier 2 on the development of disruptive behavior, compared to Tier 1. The results of the study reveal the qualitative impact of the assignment to Tier 2 on disruptive behavior. Through two different multi-level analyses and the resulting models, no significant difference (regarding the mean of disruptive behavior) between Phases B and C can be determined. This implies that no significant difference in students' disruptive behavior became measurable through assignment to Tier 2. However, there is a significant difference between the baseline (Phase A) and the assignment to Tier 2 (Phase C). As this is not adjusted in Model 1, Model 2 must be calculated to provide further information about the phase difference from Phase B to C.

All the students accompanied by single-case research in this study have moved on to the second stage of this MTSS. This implies that after implementing the GBG at the universal tier, they also received a DBRC, which supported them in achieving a pre-defined goal regarding behavior. The evaluation of the second tier, which involved continuing the use of the GBG and the additional use of the DBRC for these students, concluded that the two measures can be used together, but there is a slight further reduction of −0.02 in disruptive behavior (trend effect of Model 2). The behavior gradually subsides, but this effect cannot be attributed to Tier 2. If a level effect were measurable, it would indicate a special effectiveness of the DBRC in combination with the GBG; however, this cannot be inferred from the results. Follow-up studies should examine this point and apply a methodological design, which is often difficult to implement in practice, to determine the isolated effects of the DBRC. Nevertheless, the overall concept of Tiers 1 and 2 is effective in reducing disruptive behavior.

An additional significant point is the definition of the variable and the resulting observation. Disruptive behavior in the classroom is one of the main factors that can be reduced by the GBG. However, the DBRC is an extremely flexible intervention. The teacher formulates a goal that should serve as a behavioral goal for the student. Thus, behavior goals can focus on other aspects of the behavior that are not captured by the variables and definition used here, resulting in a behavior change that diverts attention from disruptive behavior in the classroom.

In summary, disruptive behavior does not significantly differ between the means of Tiers 1 and 2 (level effect Model 2); however, a further decrease in behavior is evidenced through the trend effect of Model 2. The exact reasons and causes are not clearly identifiable. Nevertheless, a long-term effect of the GBG is conceivable, or a paired effect of both measures together, which, however, does not show a sudden behavioral change.

The overall findings of this study contribute to the existing research landscape on intervention research in the context of disruptive behavior and tiered support. The study extends the current literature by suggesting that the combination of the two tiers investigated can be implemented in school settings and may have varying degrees of effectiveness in reducing disruptive behavior. Methodologically, the study offers valuable insights for future studies aiming to investigate complex interventions and pedagogical systems with similar research inquiries. As demonstrated by previous single case studies in related fields, utilizing students' individual developmental trajectories and within-individual reference norms can yield reliable results in assessing the effectiveness of interventions

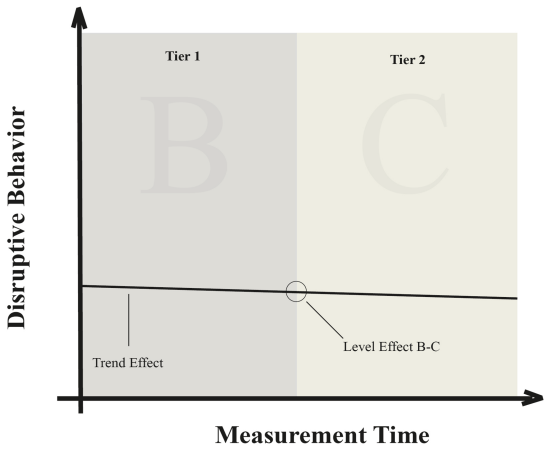

FIGURE 3
Schematic representation of Model 2.

(Franzen and Kamps, 2008; Hill and Flores, 2014; Nelson et al., 2018; Wu et al., 2019). Also the data collection method, DBR, employed in this study has proven successful in generating valid outcomes regarding disruptive behavior and teacher evaluations (Casale et al., 2017).

9.1. Limitations

Although the limitations of the study are mainly related to the research design, they are strongly determined by the pedagogical practice. The comparison of different intervention phases in a multiple baseline design is now a common procedure in intervention research (Wilbert et al., 2022). In the best case, however, the different interventions would be mixed in the internal group comparison, so that Group 1, for example, would receive the GBG first, and Group 2 would receive the DBRC first. Accordingly, Group 1 would continue with the DBRC at Tier 2, and Group 2, with the GBG. Thus, comparability would increase in terms of the effectiveness of the intervention between the different groups (Epstein and Dallery, 2022). Furthermore, taking away the intervention would be an effective method to test the combination of the two interventions and their interaction. However, from an ethical perspective, it was impossible to implement such a design. Additionally, it was impossible to change the order of the interventions in terms of the framework concept and its content. Thus, it is also impossible to make any definite statement about the sole effect of the DBRC.

Furthermore, the use of the 11-point scale could be viewed as a limitation of the study. The results show that the teachers hardly used values above six. Therefore, a 6-point Likert scale would be a suitable scale, as used in other studies (Leidig et al., 2022). Despite the comparatively low rating of the teachers, the students showed externalizing problems and disruptive behavior in the pre-survey by the ITRF.

Ultimately, it is impossible to state when the teachers made the assessment on the DBR sheet. They were initially instructed to do so after using the GBG; however, owing to the challenges of everyday practice, this response behavior may shift to a different time. These are typical problems in single-case studies and are difficult to eliminate.

One final point regarding the study's limitations is that the school routine was reinstated only a few weeks after the COVID-19 pandemic lockdown. While teachers reported no specific effects of the pandemic on the study results from their perspective, it cannot be completely ruled out.

9.2. Practical implications

The use of the GBG is effective in reducing disruptive behavior in elementary school children. The findings of this study are in line with those of previous studies (Leidig et al., 2022; Hagen et al., 2023), and emphasize the practical implementation of the GBG in the classroom. Furthermore, the GBG is an effective intervention at the first tier of support in the context of the MTSS.

The second tier requires a stronger focus on problem behavior, which the DBRC can partially achieve. However, the results of the second tier may also be related to the research design. Disruptive behavior continues to decline, but no significant difference is observed between the two tiers. Given the problems with the research design, the results must be interpreted with extreme caution.

Therefore, as a practical implication, the implementation of the GBG at the first tier and the combination of the DBRC and GBG interventions at the second tier can gradually reduce the disruptive behavior of students at inclusive elementary schools. For the practical implementation of an MTSS, it can be inferred from the results that the two interventions GBG and DBRC can be implemented together in both tiers. Additionally, it is relevant to note that the sample comprises children with disruptive behavior. When designing an MTSS with the goal of reducing externalizing or disruptive behavior, the combination of these two tiers seems to be effective and feasible.

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 Ethikkommission der Humanwissenschaftlichen Fakultät der Universität zu Köln. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

JN: original draft and visualization. JN, ToH, and JK: methodology, software, and formal analysis. JN, ToH, JK, LV, and KE: writing and editing. ThH and CH: supervision and project coordination. JN, ToH,

JK, LV, KE, CH, and ThH: conceptualization, review, and editing. All authors contributed to the article and approved the submitted version.

Funding

University funds were used for this project. There was no external funding but the authors acknowledge support for the Article Processing Charge from the DFG (German Research Foundation, 491454339).

Acknowledgments

For further collegial support the authors thank Jürgen Wilbert, Robert J. Volpe, and Julian Börger.

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

References

Adamson, R. M., Mckenna, J. W., and Mitchell, B. (2019). Supporting all students: creating a tiered continuum of behavior support at the classroom level to enhance schoolwide multi-tiered systems of support. *Prev. School Fail. Alternat. Educ. Child. Youth* 63, 62–67. doi: 10.1080/1045988X.2018.1501654

Agyapong, B., Obuobi-Donkor, G., Burback, L., and Wei, Y. (2022). Stress, burnout, anxiety and depression among teachers: a scoping review. *Int. J. Environ. Res. Public Health* 19, 1–42. doi: 10.3390/ijerph191710706

Alresheed, F., Hott, B. L., and Bano, C. (2013). Single subject research: a synthesis of analytic methods. *J. Spec. Educ. Apprenticeship* 2, 1–18.

Barkmann, C., and Schulte-Markwort, M. (2012). Prevalence of emotional and behavioural disorders in German children and adolescents: a meta-analysis. *J. Epidemiol. Community Health* 66, 194–203. doi: 10.1136/jech.2009.102467

Barrish, H. H., Saunders, M., and Wolf, M. M. (1969). Good behavior game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. *J. Appl. Behav. Anal.* 2, 119–124. doi: 10.1901/jaba.1969.2-119

Barth, J. M., Dunlap, S. T., Dane, H., Lochman, J. E., and Wells, K. C. (2004). Classroom environment influences on aggression, peer relations, and academic focus. *J. Sch. Psychol.* 42, 115–133. doi: 10.1016/j.jsp.2003.11.004

Batsche, G. (2014). "Multi-tiered system of supports for inclusive schools" in *Handbook of effective inclusive schools*

Behr, J., Leidig, T., Krull, J., Spilles, M., and Hennemann, T. (2020). Multiplikatorenkonzepte zur Professionalisierung von Lehrkräften – ein systematisches Review empirischer Studien. *Lehrerbildung auf dem Prüfstand* 13, 151–176.

Boldrini, T., Ghiandoni, V., Mancinelli, E., Salcuni, S., and Solmi, M. (2023). Systematic review and Meta-analysis: psychosocial treatments for disruptive behavior symptoms and disorders in adolescence. *J. Am. Acad. Child Adolesc. Psychiatry* 62, 169–189. doi: 10.1016/j.jaac.2022.05.002

Bradshaw, C. P., Schaeffer, C. M., Petras, H., and Ialongo, N. (2010). Predicting negative life outcomes from early aggressive-disruptive behavior trajectories: gender differences in maladaptation across life domains. *J. Youth Adolesc.* 39, 953–966. doi: 10.1007/s10964-009-9442-8

Briesch, A. M., Chafouleas, S. M., Nissen, K., and Long, S. (2019). A review of state-level procedural guidance for implementing multitiered Systems of Support for behavior (MTSS-B). *J. Posit. Behav. Interv.* 22, 131–144. doi: 10.1177/109830071988470

Casale, G., Grosche, M., Volpe, R. J., and Hennemann, T. (2017). Zuverlässigkeit von Verhaltensverlaufsdiagnostik über Rater und Messzeitpunkte bei Schülern mit externalisierenden Verhaltensproblemen. *Empirische Sonderpädagogik* 9, 143–164. doi: 10.25656/01:15010

Casale, G., Volpe, R. J., Daniels, B., Hennemann, T., Briesch, A. M., and Grosche, M. (2018). Measurement invariance of a universal behavioral screener across samples from the USA and Germany. *Eur. J. Psychol. Assess.* 34, 87–100. doi: 10.1027/1015-5759/a000447

Chaffee, R. K., Briesch, A. M., Johnson, A. H., and Volpe, R. J. (2017). A Meta-analysis of class-wide interventions for supporting student behavior. *Sch. Psychol. Rev.* 46, 149–164. doi: 10.17105/SPR-2017-0015.V46-2

Charlton, C. T., Moulton, S., Sabey, C. V., and West, R. (2021). A systematic review of the effects of schoolwide intervention programs on student and teacher perceptions of school climate. *J. Posit. Behav. Interv.* 23, 185–200. doi: 10.1177/1098300720940168

Cohen, M. A. (1998). The monetary value of saving a high-risk youth. *J. Quant. Criminol.* 14, 5–33. doi: 10.1023/A:1023092324459

Cui, L. (2022). The role of teacher-student relationships in predicting Teachers' occupational wellbeing, emotional exhaustion, and enthusiasm. *Front. Psychol.* 13:896813. doi: 10.3389/fpsyg.2022.896813

Daniels, B., Volpe, R. J., Briesch, A. M., and Fabiano, G. A. (2014). Development of a problem-focused behavioral screener linked to evidence-based intervention. *Sch. Psychol. Q.* 29, 438–451. doi: 10.1037/spq0000100

Donaldson, J. M., Fisher, A. B., and Kahng, S. (2017). Effects of the good behavior game on individual student behavior. *Behav. Anal. Res. Pract.* 17, 207–216. doi: 10.1037/bar0000016

Eagle, J. W., Dowd-Eagle, S. E., Snyder, A., and Holtzman, E. G. (2015). Implementing a multi-tiered system of support (MTSS): collaboration between school psychologists and administrators to promote systems-level change. *J. Educ. Psychol. Consult.* 25, 160–177. doi: 10.1080/10474412.2014.929960

Epstein, L. H., and Dallery, J. (2022). The family of single-case experimental designs. *Harvard Data Sci. Rev.* 4. doi: 10.1162/99608f92.ff9300a8

Fergusson, D. M., Horwood, L. J., and Ridder, E. M. (2005). Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood. *J. Child Psychol. Psychiatry* 46, 837–849. doi: 10.1111/j.1469-7610.2004.00387.x

Flower, A., Mckenna, J. W., Bunuan, R. L., Muething, C. S., and Vega, R. (2014). Effects of the good behavior game on challenging behaviors in school settings. *Rev. Educ. Res.* 84, 546–571. doi: 10.3102/0034654314536781

Fossum, S., Handegard, B. H., Martinussen, M., and Morch, W. T. (2008). Psychosocial interventions for disruptive and aggressive behaviour in children and adolescents: a meta-analysis. *Eur. Child Adolesc. Psychiatry* 17, 438–451. doi: 10.1007/s00787-008-0686-8

Fox, R. A., Leif, E. S., Moore, D. W., Furlonger, B., Anderson, A., and Sharma, U. (2021). A systematic review of the facilitators and barriers to the sustained implementation of school-wide positive behavioral interventions and supports. *Educ. Treat. Child.* 45, 105–126. doi: 10.1007/s43494-021-00056-0

Franzen, K., and Kamps, D. (2008). The utilization and effects of positive behavior support strategies on an Urban School playground. *J. Posit. Behav. Interv.* 10, 150–161. doi: 10.1177/1098300708316260

Gage, N. A., Lee, A., Grasley-Boy, N., and George, H. P. (2018). The impact of school-wide positive behavior interventions and supports on school suspensions: a statewide quasi-experimental analysis. *J. Posit. Behav. Interv.* 20, 217–226. doi: 10.1177/1098300718768204

Gage, N. A., Sugai, G., Lewis, T. J., and Brzozowy, S. (2015). Academic achievement and school-wide positive behavior supports. *J. Disabil. Policy Stud.* 25, 199–209. doi: 10.1177/1044207313505647

Gagnon, J. C., Ruiz, E., Mathur, S. R., and Mason-Williams, L. (2022). Interventions addressing incarcerated youth behavior: a review of literature. *Am. J. Orthopsychiatry* 92, 405–417. doi: 10.1037/ort0000621

Groves, E. A., and Austin, J. L. (2017). An evaluation of interdependent and independent group contingencies during the good behavior game. *J. Appl. Behav. Anal.* 50, 552–566. doi: 10.1002/jaba.393

Groves, E. A., and Austin, J. L. (2019). Does the good behavior game evoke negative peer pressure? Analyses in primary and secondary classrooms. *J. Appl. Behav. Anal.* 52, 3–16. doi: 10.1002/jaba.513

Hagen, T., Nitz, J., Brack, F., Hövel, D. C., and Hennemann, T. (2023). Effekte des Good Behavior Game bei Grundschüler_innen mit externalisierenden Verhaltensproblemen. *Lernen und Lernstörungen*, 1–14. doi: 10.1024/2235-0977/a000410

Hanisch, C., Casale, G., Volpe, R. J., Briesch, A. M., Richard, S., Meyer, H., et al. (2019). Gestufte Förderung in der Grundschule. *Prävent. Gesundheitsförderung* 14, 237–241. doi: 10.1007/s11553-018-0700-z

Hanisch, C., Richard, S., Eichelberger, I., Greimel, L., and DöPfner, M. (2018). *Schulbasiertes Coaching bei Kindern mit expansivem Problemverhalten (SCEP)*. Göttingen: Hogrefe Verlag.

Herman, K. C., Hickmon-Rosa, J. E., and Reinke, W. M. (2017). Empirically derived profiles of teacher stress, burnout, self-efficacy, and coping and associated student outcomes. *J. Posit. Behav. Interv.* 20, 90–100. doi: 10.1177/1098300717732066

Hill, D. A., and Flores, M. M. (2014). Modeling positive behavior interventions and supports for preservice teachers. *J. Posit. Behav. Interv.* 16, 93–101. doi: 10.1177/1098300713478665

Horn, E. P., Mccormick, M. P., O'connor, E. E., Mcclowry, S. G., and Hogan, F. C. (2021). Trajectories of teacher–child relationships across kindergarten and first grade: the influence of gender and disruptive behavior. *Early Child. Res. Q.* 55, 107–118. doi: 10.1016/j.ecresq.2020.10.003

Huber, C. G. M. (2013). "Inklusive Schulentwicklung durch response-to- intervention (RTI) – Realisierungsmöglichkeiten des RTI-Konzepts im Förderbereich Lesen" in *Gemeinsam Leben* (Weinheim: Beltz Juventa), 79–90.

Huber, C., and Rietz, C. (2015). Direct Behavior Rating (DBR) als Methode zur Verhaltensverlaufsdiagnostik in der Schule: Ein systematisches Review von Methodenstudien. *Empirische Sonderpädagogik* 7, 75–98. doi: 10.25656/01:10819

Husky, M. M., Boyd, A., Bitfoi, A., Carta, M. G., Chan-Chee, C., Goelitz, D., et al. (2018). Self-reported mental health in children ages 6-12 years across eight European countries. *Eur. Child Adolesc. Psychiatry* 27, 785–795. doi: 10.1007/s00787-017-1073-0

Institute of Medicine Committee on Prevention of Mental Disorders (1994) in *Reducing risks for mental disorders: Frontiers for preventive intervention research*. eds. P. J. Mrazek and R. J. Haggerty (Washington, DC: National Academies Press)

Joslyn, P. R., Donaldson, J. M., Austin, J. L., and Vollmer, T. R. (2019). The good behavior game: a brief review. *J. Appl. Behav. Anal.* 52, 811–815. doi: 10.1002/jaba.572

Joslyn, P. R., and Vollmer, T. R. (2020). Efficacy of teacher-implemented good behavior game despite low treatment integrity. *J. Appl. Behav. Anal.* 53, 465–474. doi: 10.1002/jaba.614

Joslyn, P. R., Vollmer, T. R., and Hernández, V. (2014). Implementation of the good behavior game in classrooms for children with delinquent behavior. *Acta de Investigación Psicológica* 4, 1673–1682. doi: 10.1016/S2007-4719(14)70973-1

Klasen, F., Meyrose, A. K., Otto, C., Reiss, F., and Ravens-Sieberer, U. (2017). Psychische Auffälligkeiten von Kindern und Jugendlichen in Deutschland. *Monatsschr. Kinderheilkd.* 165, 402–407. doi: 10.1007/s00112-017-0270-8

Klipker, K., Baumgarten, F., Göbel, K., Lampert, T., and Hölling, H. (2018). Psychische Auffälligkeiten bei Kindern und Jugendlichen in Deutschland – Querschnittergebnisse aus KiGGS Welle 2 und Trends. *J. Health Monit.* 3, 37–45. doi: 10.17886/RKI-GBE-2018-077

Kokkinos, C. M., Panayiotou, G., and Davazoglou, A. M. (2005). Correlates of teacher appraisals of student behaviors. *Psychol. Sch.* 42, 79–89. doi: 10.1002/pits.20031

Krause, A., Goldberg, B., D'agostino, B., Klan, A., Rogers, M., Smith, J. D., et al. (2021). The association between problematic school behaviours and social and emotional development in children seeking mental health treatment. *Emot. Behav. Diffic.* 25, 278–290. doi: 10.1080/13632752.2020.1861852

Lastrapes, R. E. (2014). Using the good behavior game in an inclusive classroom. *Interv. Sch. Clin.* 49, 225–229. doi: 10.1177/1053451213509491

Lee, A., and Gage, N. A. (2020). Updating and expanding systematic reviews and meta-analyses on the effects of school-wide positive behavior interventions and supports. *Psychol. Sch.* 57, 783–804. doi: 10.1002/pits.22336

Leidig, T., Casale, G., Wilbert, J., Hennemann, T., Volpe, R. J., Briesch, A., et al. (2022). Individual, generalized, and moderated effects of the good behavior game on at-risk primary school students: a multilevel multiple baseline study using behavioral progress monitoring. *Front. Educ.* 7:917138. doi: 10.3389/feduc.2022.917138

Liaupsin, C. J., and Scott, T. M. (2007). "Disruptive behavior" in *Evidence-based interventions for students with learning and behavioral challenges*. eds. R. J. Morris and N. Mather (New York: Routledge).

Lipien, L., Kirby, M., and Ferron, J. M. (2023). "Single-case designs" in *Handbook of applied behavior analysis*, 347–362.

Loeber, R., Burke, J. D., and Pardini, D. A. (2009). Development and etiology of disruptive and delinquent behavior. *Annu. Rev. Clin. Psychol.* 5, 291–310. doi: 10.1146/annurev.clinpsy.032408.153631

Lopez Jimenez, J., Valero-Valenzuela, A., Anguera, M. T., and Diaz Suarez, A. (2016). Disruptive behavior among elementary students in physical education. *Springerplus* 5:1154. doi: 10.1186/s40064-016-2764-6

Majeika, C. E., Wilkinson, S., and Kumm, S. (2020). Supporting student behavior through behavioral contracting. *Teach. Except. Child.* 53, 132–139. doi: 10.1177/0040059920952475

Nelen, M. J. M., Willemse, T. M., Van Oudheusden, M. A., and Goei, S. L. (2020). Cultural challenges in adapting SWPBIS to a Dutch context. *J. Posit. Behav. Interv.* 22, 105–115. doi: 10.1177/1098300719876096

Nelson, M. A., Caldarella, P., Hansen, B. D., Graham, M. A., Williams, L., and Wills, H. P. (2018). Improving student behavior in art classrooms: an exploratory study of CW-FIT tier 1. *J. Posit. Behav. Interv.* 20, 227–238. doi: 10.1177/1098300718762744

O'connor, K. M., and Hayes, B. (2020). How effective are targeted interventions for externalizing behavior when delivered in primary schools? *Int. J. School Educ. Psychol.* 8, 161–173. doi: 10.1080/21683603.2018.1530157

Ozamiz-Etxebarria, N., Idoiaga Mondragon, N., Bueno-Notivol, J., Perez-Moreno, M., and Santabarbara, J. (2021). Prevalence of anxiety, depression, and stress among teachers during the COVID-19 pandemic: a rapid systematic review with Meta-analysis. *Brain Sci.* 11, 1–14. doi: 10.3390/brainsci11091172

Parker, R. I., and Vannest, K. (2009). An improved effect size for single-case research: nonoverlap of all pairs. *Behav. Ther.* 40, 357–367. doi: 10.1016/j.beth.2008.10.006

Pearce, J., Mann, M. K., Jones, C., Van Buschbach, S., Olff, M., and Bisson, J. I. (2012). The most effective way of delivering a train-the-trainers program: a systematic review. *J. Contin. Educ. Heal. Prof.* 32, 215–226. doi: 10.1002/chp.21148

Pennington, B., and Mccomas, J. J. (2017). Effects of the good behavior game across classroom contexts. *J. Appl. Behav. Anal.* 50, 176–180. doi: 10.1002/jaba.357

Pines, A. M., and Keinan, G. (2005). Stress and burnout: the significant difference. *Personal. Individ. Differ.* 39, 625–635. doi: 10.1016/j.paid.2005.02.009

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., and Rohde, L. A. (2015). Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J. Child Psychol. Psychiatry* 56, 345–365. doi: 10.1111/jcpp.12381

Posit-Team (2022). *RStudio: Integrated development environment for R*. Boston: PBC, Posit Software.

Pyle, K., and Fabiano, G. A. (2017). Daily report card intervention and attention deficit hyperactivity disorder: a Meta-analysis of single-case studies. *Except. Child.* 83, 378–395. doi: 10.1177/0014402917706370

Ravens-Sieberer, U., Kaman, A., Devine, J., Loffler, C., Reiss, F., Napp, A. K., et al. (2022). The mental health and health-related behavior of children and parents during the COVID-19 pandemic. *Dtsch. Arztebl. Int.* 119, 436–437. doi: 10.3238/arztebl.m2022.0173

Reaves, S., Mcmahon, S. D., Duffy, S. N., and Ruiz, L. (2018). The test of time: a meta-analytic review of the relation between school climate and problem behavior. *Aggress. Violent Behav.* 39, 100–108. doi: 10.1016/j.avb.2018.01.006

Reed, D. F., and Kirkpatrick, C. Metropolitan Educational research consortium, R.V.A. (1998). *Disruptive students in the classroom: A review of the literature*. V. M. Publications.

Reynolds, M. C. (1962). A framework for considering some issues in special education. *Except. Child.* 28, 367–370. doi: 10.1177/001440296202800705

Rivenbark, J. G., Odgers, C. L., Caspi, A., Harrington, H., Hogan, S., Houts, R. M., et al. (2018). The high societal costs of childhood conduct problems: evidence from administrative records up to age 38 in a longitudinal birth cohort. *J. Child Psychol. Psychiatry* 59, 703–710. doi: 10.1111/jcpp.12850

Ruiz-Olivares, R., Pino, M. J., and Herruzo, J. (2010). Reduction of disruptive behaviors using an intervention based on the good behavior game and the say-do-report correspondence. *Psychol. Sch.* 47, 1046–1058. doi: 10.1002/pits.20523

Schmerse, D., and Zitzmann, S. (2021). Early school adjustment: do social integration and persistence mediate the effects of school-entry skills on later achievement? *Learn. Instr.* 71:101374. doi: 10.1016/j.learninstruc.2020.101374

Serrano, A., Sanz, R., Cabanillas, J. L., and Lopez-Lujan, E. (2023). Socio-emotional competencies required by school counsellors to manage disruptive Behaviours in secondary schools. *Children (Basel)* 10:10. doi: 10.3390/children10020231

Simonsen, B., Robbie, K., Meyer, K., Freeman, J., Everett, S., and Feinberg, A. (2021). *Multi-tiered System of Supports (MTSS) in the Classroom*. Center on PBIS, University of Oregon.

Skinner, A. T., Babinski, L. M., and Gifford, E. J. (2014). Teachers' expectations and self-efficacy for working with bullies and victims. *Psychol. Sch.* 51, 72–84. doi: 10.1002/pits.21735

Steed, E. A., Noh, J., and Heo, K. H. (2014). A cross-cultural comparison of positive behavioral interventions and supports in early childhood classrooms in the United States and South Korea. *Infants Young Child.* 27, 30–42. doi: 10.1097/IYC.0b013e3182a4ec46

Stoiber, K. C., and Gettinger, M. (2016). "Multi-tiered Systems of Support and Evidence-Based Practices" in *Handbook of response to intervention: The science and practice of multi-tiered Systems of Support*. eds. S. R. Jimerson, M. K. Burns and A. M. Vanderheyden (Boston, MA: Springer), 121–141.

Swoboda, C. M., Kratochwill, T. R., and Levin, J. R. (2010). Conservative dual-criterion method for single-case research: A guide for visual analysis of AB, ABAB, and multiple-baseline designs. University of Wisconsin–Madison, Wisconsin Center for Education Research.

Taylor, J. C., and Hill, D. (2017). Using daily behavior report cards during extended school year Services for Young Students with intellectual and developmental disabilities. *Educ. Treat. Child.* 40, 525–546. doi: 10.1353/etc.2017.0023

Tehrani, H. D., Yamini, S., and Vazsonyi, A. T. (2023). The effectiveness of parenting program components on disruptive and delinquent behaviors during early and middle childhood: a component network meta-analysis. *J. Exp. Criminol.* doi: 10.1007/s11292-023-09562-0

Vannest, K. J., Davis, J. L., Davis, C. R., Mason, B. A., and Burke, M. D. (2010). Effective intervention for behavior with a daily behavior report card: a Meta-analysis. *Sch. Psychol. Rev.* 39, 654–672. doi: 10.1080/02796015.2010.12087748

Volpe, R. J., Briesch, A. M., and Vanderheyden, A. (2012). Generalizability and dependability of single-item and multiple-item direct behavior rating scales for engagement and disruptive behavior. *Sch. Psychol. Rev.* 41, 246–261. doi: 10.1080/02796015.2012.12087506

Volpe, R. J., Fabiano, G. A., and Pelham, W. E. Jr. (2013). *Daily behavior report cards: An evidence-based system of assessment and intervention*. New York: Guilford Publications.

Volpe, R. J., Yeung, T. S., Casale, G., Krull, J., Briesch, A. M., and Henneman, T. (2020). Evaluation of a German language school-based universal screening for student social, emotional, and behavioral risk. *Int. J. School Educ. Psychol.* 9, S10–S20. doi: 10.1080/21683603.2020.1717699

Weist, M. D., Eber, L., Horner, R., Splett, J., Putnam, R., Barrett, S., et al. (2018). Improving multitiered Systems of Support for students with "internalizing" emotional/behavioral problems. *J. Posit. Behav. Interv.* 20, 172–184. doi: 10.1177/1098300717753832

Whitten, L. (2009). Intervention for disruptive children shows long-term benefits. *Nida Notes* 22, 4–6.

Wilbert, J., Börnert-Ringleb, M., and Lüke, T. (2022). Statistical power of piecewise regression analyses of single-case experimental studies addressing behavior problems. *Front. Educ.* 7, 1–13. doi: 10.3389/feduc.2022.917944

Wilbert, J., and Lueke, T. (2022). Scan: Single-case data analyses for single and multiple baseline designs.

Wu, Y. C., Chen, P. Y., Tsai, S. P., Tsai, S. F., Chou, Y. C., and Chiu, C. Y. (2019). The effects of the class-wide function-related intervention teams (CW-FIT) on behaviors of an elementary student with autism spectrum disorder (ASD) in an inclusive classroom in Taiwan. *Int. J. Dev. Disabil.* 65, 368–377. doi: 10.1080/20473869.2019.1647031

Yamasaki, K., and Nishida, N. (2009). The relationship between three types of aggression and peer relations in elementary school children. *Int. J. Psychol.* 44, 179–186. doi: 10.1080/00207590701656770

Zuniga, A. N., and Cividini-Motta, C. (2021). Using class pass intervention to decrease disruptive behavior in young children. *J. Behav. Educ.* 31, 524–542. doi: 10.1007/s10864-020-09411-4