Universiteit Leiden
The Netherlands

School-wide positive behavioral interventions and supports in Dutch elementary schools: exploring effects
Nelen, M.J.M.; Scholte, R.H.J.; Blonk, A.M.; Veld, W.M. van der; Nelen, W.B.L.; Denessen, E.J.P.G.

RESEARCH ARTICLE

WILEY

School-wide positive behavioral interventions and supports in Dutch elementary schools: Exploring effects

Monique J. M. Nelen[1,2] | Ron H. J. Scholte[2,3] | Anita M. Blonk[1] | William M. van der Veld[2] | Wendy B. L. Nelen[4] | Eddie Denessen[2,5]

[1]Windesheim University of Applied Sciences, Zwolle, The Netherlands

[2]Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

[3]Tranzo, Tilburg University, Tilburg, The Netherlands

[4]Praktikon, Nijmegen, The Netherlands

[5]Education and Child Studies, Leiden University, The Netherlands

Correspondence
Monique J. M. Nelen, Windesheim University of Applied Sciences, Campus 2-6, 8000 GB Zwolle, The Netherlands.
Email: mjm.nelen@windesheim.nl

Abstract

In 2009, School-Wide Positive Behavioral Interventions and Supports (SWPBIS) was introduced in the Netherlands to support schools in creating safe learning environments. In this longitudinal study, we explored effects of SWPBIS on student outcomes in the Netherlands. Fidelity of implementation of SWPBIS has been associated with improved student outcomes. The purpose of this study was to examine the relation between changes in fidelity and student outcomes. A total of 66 elementary schools ($n = 14{,}256$ students) were followed for 3 years (2015–2018). We collected yearly data on fidelity, social safety (consisting of students' social well-being, general feeling of safety, harassment, prevalence of unsafe locations in and around schools), behavior incidents, and additional behavioral support. Using repeated measures analysis of variances, we saw an increase in fidelity scores and a decline in the percentage of students stating there were unsafe locations in and around school. Multiple regression analyses showed that changes in fidelity were related to changes in both students' social well-being and the number of behavior incidents. Limitations were discussed, such as the absence of comparison

schools not implementing SWPBIS, and schools at different stages of implementation, and we accounted for missing data.

In 2009, School-Wide Positive Behavioral Interventions and Supports (SWPBIS) was introduced in the Netherlands to support schools in dealing with problem behavior and creating safe environments. SWPBIS was originally developed in the US in the 1980s by researchers from the University of Oregon (Sugai & Simonsen, 2012), and more than 26,000 U.S. schools are currently working with SWPBIS. Its aim is to develop school-wide systems and procedures that promote positive changes in student behavior by targeting staff behavior (Bradshaw et al., 2010). SWPBIS is a framework, not a method with specific protocols or standardized interventions: Strategies and interventions are developed and modified in alignment with the context of the individual school, referred to as contextual fit (McIntosh et al., 2010). Research has shown that SWPBIS resulted in a decrease in problem behavior, an increase in prosocial skills and perceptions of school safety, and an improvement of the overall school climate (e.g., Bradshaw et al., 2008, 2009, 2012; Horner et al., 2009, 2010; Waasdorp et al., 2012). Most SWPBIS research has been U.S.-oriented, although other countries such as Norway and Australia have been building evidence for the effectiveness of SWPBIS as well (Sørlie & Ogden, 2015; Yeung et al., 2016). Implementing SWPBIS with fidelity has been shown to be important for achieving positive outcomes (McIntosh et al., 2013). In this study, we aimed to explore effects of SWPBIS in the Netherlands, with particular attention to the role of fidelity of implementation. We followed 66 elementary schools (14,256 students) for 3 years, collecting data on fidelity of implementation and student outcomes.

1 | SWPBIS FEATURES

Sugai and Horner (2009) described the theoretical and conceptual characteristics of SWPBIS as (a) the behavioral foundation of SWPBIS, (b) emphasis on prevention in a multitiered system of behavior support, (c) teaching of behavior, (d) the use of evidence-based or research-based practices, (e) the implementation of systems that support effective practices related to school safety, and (f) the on-going collection and use of behavioral data to develop (preventive) strategies. The multitiered system of student support (Greenwood et al., 2008) contains universal interventions for all students (Tier 1), targeted interventions for students who need additional support (Tier 2), and individual interventions for students with chronic or severe behavioral needs who need individual support (Tier 3). At Tier 1, a SWPBIS school typically has established school-wide expectations (such as "Be responsible") that are being taught, systematically acknowledges positive student behavior, and has a system for handling problem behavior, including procedures for how to respond to problem behavior with consistent consequences (Office of Special Education Programs OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports, 2015). Data-driven decision making is a central feature of SWPBIS (McIntosh et al., 2018). Behavioral data such as office discipline referrals (ODRs) are collected and used to develop and evaluate preventive interventions. Systems change and research-validated practices are used to reach valued outcomes that are defined and operationalized by the school (Sugai, O'Keeffe, & Fallon, 2012). A SWPBIS leadership team (a representative group of stakeholders including educators, school administrator(s), family members, and students) is responsible for the implementation process at the school, establishing local capacity and expertise, setting up majority agreements and commitments, measuring fidelity of implementation, and outcome evaluation (Lewis et al., 2016; Sailor et al., 2009). In the United States, school-based leadership teams receive further support from district- and state-level leadership teams (Office of Special Education Programs OSEP Technical Assistance Center on Positive Behavioral Interventions and Supports).

All the separate components mentioned above are part of the SWPBIS framework and draw from several decades of systematic research in education, mental health, and behavior analysis (Horner et al., 2010). The efficacy of SWPBIS is based on focusing on the whole school approach, emphasizing the multiple tiers of support that are delivered as early as possible, tying educational practices to organizational systems needed to deliver these practices with fidelity, and the systematic use of data for decision making (Center on Positive Behavioral Interventions and Supports, 2015). Adapting the framework to the school context is crucial for successful implementation (McIntosh et al., 2010). This not only applies to implementation of SWPBIS in diverse U.S. cultural contexts, but also to implementation in other countries (Nelen et al., 2020a). However, adaptations made to make SWPBIS fit more closely to the (national) school context must be in line with the conceptual foundations of the framework to avoid weakening the efficacy (Smith et al., 2011). When SWPBIS was introduced in the Netherlands, essential features of the framework were formulated in recognizable and culturally acceptable words, and interventions and strategies were adjusted to fit the Dutch schools. Nelen et al. (2020a) have described the process of cultural adaptation of the framework to the Dutch educational context.

2 | SWPBIS IN THE NETHERLANDS

Discussing effects of SWPBIS in a country requires understanding of the cultural context. In the 2015–2016 school year, there were 6431 elementary schools (grades 1–8, ages 4–12 years) in the Netherlands. Many elementary schools are relatively small (50% of all elementary schools have fewer than 200 students, $M = 224$ students). The average class size in elementary school is approximately 24 students. Almost all schools are funded by the Dutch government, as long as prescriptive goals are achieved. Dutch schools are known for their high (teacher) autonomy (Organisation for Economic Co-operation and Development OECD, 2011). Every school is free to choose its curriculum and methods, achievement measures, and staff-to-student ratio. A national inspectorate monitors the quality of education in the schools. Parents are free to choose a school, and costs are minimal.

In 2009, a consortium of universities of applied sciences and youth care agencies introduced SWPBIS in the Netherlands and initiated PBIS coach training. The consortium presented several adaptations to SWPBIS procedures. As problem behavior is mostly classroom-managed and ODRs do not exist in Dutch schools, a behavior incident form was developed for the ongoing use of behavioral data. Collecting behavior incident data for preventive reasons is not common in Dutch schools. Therefore, during SWPBIS implementation, schools are usually coached on determining when, what and how to register. In 2014, a Dutch version of the Schoolwide Information System (SWIS; May et al., 2010) was introduced in the Netherlands. As openly praising students in the Netherlands is often considered "over the top", the introduction of token economy systems initially met with some resistance from teachers (Nelen et al., 2020a). However, research on the use of fidelity measures in Dutch schools showed that feedback and acknowledgement for positive student behavior was fully implemented at most schools (Nelen et al., 2020b). This suggests that culturally appropriate ways of reinforcing student behavior were found (such as group awards or "thumbs up"). Finally, culturally adaptive ways of coaching were developed, taking into account the high degree of autonomy of Dutch teachers. There are currently different modalities for supporting schools in implementing SWPBIS in the Netherlands: schools can be coached by a SWPBIS coach, networks of SWPBIS schools have arisen, and some schools have started SWPBIS without the guidance of a coach (Nelen et al., 2020b). Today, SWPBIS has been implemented in approximately 350 schools (approximately 4.5% of all Dutch schools), mostly elementary schools.

3 | FIDELITY OF IMPLEMENTATION

Many studies have reported that implementing SWPBIS with fidelity is associated with positive school outcomes such as improvement of school climate and safety, and a decrease in behavioral problems (e.g., Bradshaw et al., 2009; Simonsen et al., 2012). Fidelity of implementation is the extent to which components of an intervention, as conceptualized in a

theoretical model or manual, are implemented as intended (Schulte et al., 2009). In SWPBIS studies, fidelity has been operationalized by measuring to what extent the core features and standard procedures of SWPBIS were present in schools. Fidelity measures reflect core features and standard procedures and contain items on the SWPBIS leadership team (composition, procedures and universal screening), implementation (teaching behavioral expectations, problem be-havior definitions, classroom procedures, providing students with feedback and acknowledgement, stakeholder involve-ment and professional development), and evaluation (collecting discipline data, data based decision making, measuring fidelity and annual evaluation). As the process of implementation can vary across schools in different countries, measuring fidelity provides information regarding the extent to which a school has succeeded in implementing core features and procedures (McIntosh et al., 2017).

Fidelity does not happen automatically: schools work hard to contextualize and implement core features and pro-cedures. Usually, SWPBIS coaches support schools in their implementation efforts. Fixsen et al. (2009) distinguished several stages of implementation: creating readiness, initial implementation and institutionalization. Nese et al. (2019) found that most schools reached adequate implementation at Tier 1 during their second year of implementation following training. The initial years of implementation are crucial as threats like administrator or team turnover can easily lead to abandoning SWPBIS. To embed SWPBIS practices into school routines may even take three to 5 years (Sugai et al., 2008). Reaching implementation early is a strong predictor of sustained implementation (McIntosh et al., 2015).

To measure fidelity of Tier 1 implementation, several instruments have been developed. The most recently developed is the Tiered Fidelity Inventory (TFI; McIntosh et al., 2017). The SWPBIS leadership team of a school completes a questionnaire, preferably with guidance by a SWPBIS coach to ensure as much objectivity as possible. The School-wide Evaluation Tool (SET; Horner et al., 2004) is another fidelity measure, mostly used in research studies because it is considered to be a more objective measure, as it is completed by an external assessor. Both instruments are valid and reliable, and assess the same construct (Mercer et al., 2017). They both result in a total score, indicating the level at which features are realized. Higher scores mean greater fidelity. When the total score meets or exceeds a criterion (e.g., 80% for the SET and 70% for the TFI), it indicates that a school is implementing SWPBIS "with fidelity" (Mercer et al., 2017). In the present study, we used both the TFI and the SET to measure fidelity of Tier 1 implementation in Dutch schools. The TFI was chosen because it is the most recently developed and up-to-date instrument, it is brief, and it is based on the factors and features in existing validated fidelity measures. The SET was chosen to compare TFI measurements with more objective data (Nelen, 2020b).

4 | SCHOOL SAFETY

SWPBIS, when implemented with fidelity, is expected to promote safe schools, not only by reducing problem behavior or improving school climate (Horner et al., 2009), but also by enhancing schools' organizational context (Bradshaw et al., 2009). Safe schools are pivotal for learning. According to Cohen et al. (2009), positive school climate is associated with and predictive of academic achievement, school success, effective violence prevention, students' healthy development, and teacher retention. Nijs et al. (2014) stated that school environment is an important determinant of psychosocial function and may also be related to mental health. Kutsyuruba et al. (2015) found that school climate, feelings of school attachment/connectedness and personal safety are some of the most important variables for understanding school safety.

In the Netherlands, school safety is emphasized as social safety. The Dutch Ministry of Education, Culture, and Science defined three aspects of social safety: social and physical safety of students, and social well-being. When students' safety is not being violated by others, a school is considered to be safe (Nelen et al., 2018). Yearly monitoring of school safety is mandatory for Dutch schools. Although the government organizes a bi-yearly measurement of school safety, each school is free to choose an instrument for monitoring school safety. In this study, we followed the Dutch government's definition of social safety, which we operationalized as students' perceptions of school safety and the prevalence of behavior incidents. Social well-being is defined as the way students perceive their class, contacts with classmates, and being at school. Physical safety is defined as the

absence of physical harassment (such as hurting, pushing or fighting; Nelen et al., 2018). Research showing that SWPBIS contributes to improved social safety has mainly been conducted in countries outside the Netherlands (for the United States: e.g., Bradshaw et al., 2010, Canada: e.g., McIntosh et al., 2011, and Norway: e.g., Sørlie & Ogden, 2015). Therefore, we wanted to explore whether these results were replicable for the Netherlands.

5 | PURPOSE OF THE STUDY

Fidelity of implementation has been associated with positive student outcomes, such as a decrease in problem behavior and an increase in social safety. To examine this, some studies have used fidelity cut-off scores (meeting or exceeding a criterion) in their analyses (e.g., Simonsen et al., 2012). Others used fidelity as both a continuous and a dichotomous variable (e.g., Bradshaw et al., 2009). The relation between changes in fidelity and changes in student outcomes has been less examined. In the Netherlands, to our knowledge, research to study the relation between fidelity and student outcomes has not been done before. In Dutch schools, there are also different modalities for supporting SWPBIS implementation. To examine whether the core components of SWPBIS were being implemented as intended, measuring fidelity of implementation was important. Earlier research on the use of fidelity measures in the Netherlands showed that all items displayed in the TFI and SET were present in participating schools (Nelen et al., 2020b), and, therefore, these measures could be used to measure fidelity of implementation.

The number of Dutch schools implementing SWPBIS is relatively small, and in the Netherlands there is usually no (research) funding to finance the costs of implementation. Therefore, we decided to focus on elementary schools that were already implementing SWPBIS, rather than on schools that started at study onset. For 3 consecutive years, we measured fidelity of Tier 1 implementation, students' perceptions of social safety and the prevalence of behavior incidents. To determine the distribution of the multitiered model in participating Dutch schools, we also collected data on the percentage of students receiving additional support for their behavior. Our research questions were: (1) To what extent do fidelity of Tier 1 SWPBIS implementation and student outcomes (i.e., students' perceptions of social safety, the prevalence of behavior incidents, and the percentage of students receiving additional support for behavior) in Dutch elementary schools change over time? (2) What is the relation between SWPBIS Tier 1 fidelity of implementation and student outcomes in participating schools? (3) Is an increase in SWPBIS Tier 1 fidelity of implementation related to improvement in student outcomes in participating schools?

6 | METHOD

6.1 | Participating schools

Elementary schools implementing SWPBIS were recruited through invitations posted on Dutch SWPBIS websites, flyers distributed at the annual Dutch SWPBIS conference, and invitations sent by several SWPBIS expertise centers (mostly indirectly via SWPBIS coaches). Of 83 schools asked to participate in the 3-year study, 76 initially accepted the invitation. Of these schools, six schools declined before study onset. During data collection, four schools withdrew due to management changes or not being able to provide the data requested. In the end, 66 schools participated for all 3 years. Effect sizes for SWPBIS have been reported to vary across studies from relatively small ($d = 0.31$; Simonsen et al., 2012) to very large ($d = 2.63$; Bradshaw et al., 2010), and to depend on the variables assessed (Horner et al., 2009). For student outcomes, mean effect sizes are around $d = 0.32$ (Simonsen et al., 2012), and for fidelity measures effect sizes are well above 1 (d varies between 1.08 and 2.63). Based on the smallest reported effect size (0.31), an alpha of 0.05 and a power of 0.80, a total sample size of 52 schools for a repeated measures analysis of variance (ANOVA) was considered large enough to detect significant effects.

All participating schools started implementing SWPBIS before study onset. Average duration of implementation at study onset was 22.97 months ($SD = 16.53$ months, range: 2–74 months). All schools received

TABLE 1 Descriptive data for participating schools at T1 (N = 66)

	M	*Min*	*Max*	*SD*
Number of students	216	57	476	104.73
Number of teachers	17.35	6	42	8.33
Number of classes	9.29	3	19	3.97

support from a trained SWPBIS coach, mainly at the beginning of the implementation process. The training contained, among others, issues such as implementing and monitoring fidelity of SWPBIS implementation. Authors had no involvement with implementing SWPBIS in participating schools. The process of implementation was not part of this study. Schools in our sample were comparable with other Dutch elementary schools in size, location, and affiliation. Twenty-five schools reported they were located in a multiproblem neighborhood. We defined this as a neighborhood where multiple problems occur, such as unemployment, violence, criminality, addiction-related problems, and health problems such as higher mortality rate and obesity (e.g., Marlet et al., 2009). See Table 1 for summary information about numbers of teachers, students, and classes at participating schools.

7 | PROCEDURE

7.1 | Data collection

Data were collected for 3 consecutive years, with a focus on the first and last wave (T1 and T3), in repeated measurements of fidelity of Tier 1 implementation and student outcomes (social safety, behavior incidents, and the percentage of students receiving additional behavioral support). All data were collected between October 2015 and August 2018. In defining our measures, we stayed as close as possible to the daily practice in schools. We chose measures that were either part of SWPBIS (behavior incident form), or part of schools' obligation to collect data on social safety (social safety monitor).

Fidelity, social safety, and the percentage of students receiving additional support were measured yearly. Data on behavior incidents were collected several times per year in 10 periods of 4 weeks each. Data collection was synchronized each year.

For behavior incidents and students receiving additional support, we asked schools to anonymize their data before sending them per email. Most data were at the school level, except for the social safety monitor; in that, individual student data were collected. Schools were invited by email to subscribe to the safety monitor. In accordance with the official survey procedure, only students from grades 7 and 8 (10–12 years old) received a login code (more than 3500 students), so they could complete the survey anonymously. The safety monitor used in this study is one of the social safety monitors officially approved by the Dutch inspectorate of education. Since monitoring social safety is prescribed by law, no parental consent for participation of students was needed. The internal review board of the research institute approved the study (ECSW 2016-2501-369). At the beginning of the school year, each school received an overview of which data were planned to be collected when. When a school did not provide the data requested, several reminder emails were sent.

8 | MEASURES

Fidelity of implementation was measured with both the TFI Tier 1 and the SET. We focused on Tier 1, because not many schools have implemented Tiers 2 and 3 yet. The TFI Tier 1 (version 2.1) has 15 questions, divided into three subscales: "Team," "Implementation," and "Evaluation" (McIntosh et al., 2017). The SET was originally designed for

academic research and is completed by an external assessor (Horner et al., 2004). It has seven subscales, "Expectations defined," "Behavior expectations taught," "Reward system," "Violations system," "Monitoring and evaluation," "Management," and "District support." There are multiple items per subscale with a total of 28 items. For each subscale, the sum score is divided by the maximum score per scale. In both measures items can be scored 2 (*fully implemented*), 1 (*partially implemented*), or 0 (*not implemented*). The total score indicates the level at which features are realized in schools in percentages. A weighted score was used for the SET total score by adding all seven subscale scores (maximum score 1 per subscale), divided by 7 and multiplied by 100. For the TFI Tier 1 total score, the sum for the 15 items was divided by 30 (total possible score) and multiplied by 100.

The TFI Tier 1 was completed first, by discussing the 15 questions to reach consensus during a SWPBIS leadership team meeting. The meeting was guided by a SWPBIS coach, who explicitly asked for substantiation of the choices made. Before the meeting, the SWPBIS coach made some observations, and briefly interviewed both students and teachers about school values and behavioral expectations, and acknowledging students. Preferably, this SWPBIS coach also was (or had been) responsible for coaching the school during SWPBIS implementation. When the school did not have a SWPBIS coach to assist with completing the measurement, one was provided (approximately 14 times). Following that, the SET was completed within 2 weeks by a different SWPBIS professional who was not familiar with the school. This professional conducted structured interviews with the administrator, staff members and students, observed the school environment, and reviewed developed products such as school policies, SWPBIS Handbook or documents, and data systems. For example, to determine how well a school's values and accompanying behavioral expectations had been taught, the assessor studied lesson plans and asked at least 15 students and 10 staff members whether they could state the values and behavioral expectations of their school.

The TFI was completed by the same assessor every year, whereas the SET assessor varied each year. All TFI and SET assessors were familiar with SWPBIS, and were selected and trained by the first author in completing both instruments. The interrater agreement of SET assessors was moderate ($k = 0.58$) when measured in an earlier study on the use of TFI and SET in Dutch schools (Nelen et al., 2020b). That study also included data for the first fidelity measurements at T1 used in this study (66 of the 117 schools included in that study). The interrater agreement for the TFI was not calculated in that study, because scoring TFI items is based on discussions in the SWPBIS leadership team, which makes independent scoring difficult. For a more detailed description of the use of these fidelity measures in Dutch schools, see Authors 2 (2019).

Social safety was measured with an online survey measuring perceptions of social safety and required interventions, and harassment (Mooij et al., 2011). The survey consists of eight different topics. For example, "About school," "Feeling safe," "Being bullied," and "Being a bully." An example of a question was "Are you being bullied at school?" This question could be scored "Every day," "Every week, but not every day," "Sometimes, but not every week," "Almost never," or "No, never." At the beginning of each page, students were reminded that the questions were about the present school year. There was a maximum of 71 questions. Several questions are shown or hidden depending on the reaction of a previous question. Most questions were answered by multiple choice or a Likert scale. The number and content of the options varied depending on the question. In the survey, four dimensions of positive or negative aspects of social safety were distinguished: (1) the perception of safety at different school locations; (2) unacceptable behavior, represented by the prevalence of behavior incidents and substance abuse; (3) harassment of students; and (4) the perceived need for extra interventions to improve social safety in and around the school (Nelen et al., 2018). For the purpose of this study, we only used questions about students' social well-being, general feeling of safety, unsafe locations, and harassment. "Well-being" was operationalized as the average of the scores for three questions about liking one's class, number of contacts with classmates, and appreciation of these contacts (scale existing of three items, Cronbach's alpha varying from 0.61 in 2017 to 0.65 in 2016 and 2018). "General safety" was operationalized by asking students how safe they generally felt at school, on a 5-point scale (single question, validated with similar questions on safety). "Unsafe locations" was operationalized by asking students if there were various locations (total of seven, e.g., classroom, hallway, playground) in or around school where they did not feel safe at any time the past year. And "Harassment" was operationalized by asking if students had been a victim of various types of harassment at any time the past year (scale existing of six items, Cronbach's alpha

varying from 0.81 in 2016 to.97 in 2017 and 2018). Here a mean score was calculated for being bullied and/or being a victim of minor physical (e.g., hurting, pushing or fighting), social (e.g., exclusion, ignoring or threatening), material (e.g., destroying or stealing), and/or verbal (e.g., name-calling or yelling) harassment.

All data were aggregated at school level. First, answers were dichotomized (e.g., for bullying: "Almost never" and "No, never" as "0," and "Every day," "Every week, but not every day," and "Sometimes, but not every week" as "1"). Next, the answers of all students were aggregated at school level. In our example of the item on bullying, this resulted in the percentage of students who stating that they were being bullied during the last school year.

8.1 | Behavior incidents

To measure the prevalence of behavior that was not tolerated at a school, we asked schools to provide data on the number and location (in or outside class) of major and minor problem behaviors, using the schools' data collection method. Behavior was considered an incident as it interfered (or could interfere) with daily practice in schools. Minor incidents could be resolved quickly without disturbing class, with no need for support from outside class. Examples are not following a teacher's directions or name calling. Examples of major problem behaviors are physical violence, theft or vandalism. Most Dutch SWPBIS schools first define what particular behaviors can be considered as problem behavior (both minor and major), as this can vary across contexts. Second, each school decides what, when, and how to report. For this study, to support schools in collecting data on behavior incidents, we provided them with descriptions and examples based on the Dutch version of the SWIS. Data were recorded by means of the Dutch SWIS or Excel sheets, programmed by the Dutch SWPBIS consortium. For our analyses, we counted the total number of behavior incidents (major and minor incidents) and standardized this by calculating the average number of incidents per 100 students per day, for two intervals from the same 4-week period, at T1 and T3. For example: school A had 19 incidents over 18 school days in the 4-week period, and a total number of 128 students. This resulted in the following score: $[(19/18)/128] \times 100 = 0.82$ incidents per 100 students per day.

8.2 | Additional support

We asked schools to complete a form each year with the number of students receiving additional support for behavior. We defined it as extra arrangements for students, comparable to Tier 2 or Tier 3 interventions, including examples such as Check-In-Check-Out or an individual behavior plan with different rules for playing outside at recess. Each student could only be counted once. For each school, we calculated the percentage of students who received additional behavioral support.

8.3 | Analyses

In our study, the school was the unit of analysis. All analyses were performed with SPSS version 20 for Windows 10. Not all schools provided all the data requested. Therefore, the number of participating schools varied across time. We focused on the first (T1) and last (T3) waves of data collection, as we had a loss of 20% of our data if we used all three waves. Comparison between T1 and T3, 2 years apart, would allow for more change over time to occur that could be related to fidelity of implementation. We tested whether the nonresponse over time (i.e., attrition) was systematic or not. We compared the scores at T1 of schools with incomplete data at T3 with the scores at T1 of schools with complete data at T3. There were no significant differences for any of the outcome variables. We therefore concluded that the nonresponse was random and not selective, as the two groups did not differ systematically at T1.

To answer the first research question, we used within-subjects repeated measures ANOVAs to examine how group means for Tier 1 fidelity of implementation and student outcomes changed over time. To examine the relation between fidelity of implementation and student outcomes (research questions 2 and 3), we used regression analyses, as is recommended for testing associations between a predictor and outcomes. We conducted six multiple regression analyses with SET scores at T1 as the independent variable, using student outcomes at T3 as the dependent variables. We also conducted these analyses with TFI scores as the independent variable. These analyses enabled us to determine whether the level of implementation was related to changes in student outcomes (see Table 3). Next, we performed six multiple regression analyses with changes in fidelity (i.e., the difference between fidelity scores at T3 and T1) scores as the independent variable, first for SET, and second for TFI, again controlling for student outcomes at T1. These results were used to study whether changes in student outcomes depended on changes in fidelity (see Table 3). As many studies have focused on the results of schools that started implementing SWPBIS at study onset, we also calculated means for both fidelity and outcome variables for the nine schools that started in August 2015, and reported on their results separately, to give an impression of their progress across 3 years.

9 | RESULTS

Relation of outcomes and fidelity: Table 2 gives descriptive data and results of repeated measures ANOVAs, to see if student outcomes and fidelity changed over the years. Fidelity of implementation improved significantly. In addition, the percentage of students stating there were locations in or around school where they felt unsafe decreased significantly. The other variables did not change significantly, although the decrease in behavior

TABLE 2 Repeated measures ANOVAs: change over time

| | Number of schools | M_{T1} | SD_{T1} | M_{T3} | SD_{T3} | M_{T3-T1} | 95% Confidence interval of the difference | | p | Cohen's d |
							Lower	Upper		
TFI	66	57.48[a]	20.97	82.83	15.54	25.35	19.84	30.87	0.00	1.13
SET	66	68.56[a]	16.99	84.29	11.06	15.73	11.32	20.15	0.00	0.88
Well-being	39	84.38[b]	8.77	85.97	7.63	1.59	−1.53	4.71	0.31	0.17
General safety	39	85.47[c]	8.17	86.21	5.73	0.748	−2.39	3.88	0.63	0.08
Unsafe location	39	25.31[d]	10.06	20.61	9.75	−4.70	−8.46	−0.93	0.02	−0.41
Harassment	39	32.27[e]	10.03	30.18	10.14	−2.09	−6.89	2.70	0.38	−0.14
Additional support	38	4.17[f]	2.70	3.83	2.33	−0.34	−7.66	3.25	0.52	−0.13
Behavior incidents	42	1.61[g]	1.65	1.23	1.32	−0.37	−0.84	0.09	0.11	−0.25

Abbreviations: ANOVAs, analysis of variance; SET, School-wide Evaluation Tool; SWPBIS, School-Wide Positive Behavioral Interventions and Supports; TFI, Tiered Fidelity Inventory.

[a]Total score, meaning the percentage of realized SWPBIS features.

[b]The average score of liking ones class, contact with classmates, appreciation of these contacts, and liking being at school, in percentages.

[c]The percentage of students stating they generally felt safe.

[d]The percentage of students stating there were various locations in and around school where they not felt safe at any time the past year.

[e]The percentage of students stating they had been a victim of various types of harassment at any time the past year.

[f]The percentage of students receiving additional behavioral support.

[g]Incidents per 100 students per day.

incidents showed a small effect. For the nine schools that started implementing SWPBIS just before study onset, all means for student outcome variables improved, but the number of cases was too low to draw conclusions. As there was a considerable variation in months of SWPBIS implementation for participating schools that could have influenced the results found, we checked whether using months of implementation as a between-subjects factor in the repeated measures ANOVAs revealed any differences for student outcomes. This was not the case.

In Table 3, the multiple regression analyses with student outcomes at T3 as dependent variables and TFI and SET scores as independent variable are displayed. Whereas ANOVAs use group means, multiple regression analyses were conducted to identify patterns in individual school scores. On the first row of Table 3, the contribution of Well-being at T1 to predicting Well-being at T3 is presented, interpreted as the stability of Well-being scores. For all variables stability appeared to be low, although for two variables (Well-being and Behavior incidents) there were statistically significant β values. For Well-being, the β value was .34 ($p < 0.05$), indicating that stability was not perfect, so there was change in

TABLE 3 Regression analyses with SET or TFI and student outcomes at T3 as dependent variables

	Level of fidelity					Change in fidelity			
	B	β	B	β		B	β	B	β
	SET	SET	TFI	TFI		SET	SET	SET	TFI
Well-being T1	0.30*	0.34*	0.28*	0.32	Well-being T1	0.43**	0.49**	0.37*	0.42
SET T1	−0.10	−0.21			ΔSET (T3-T1)	0.08	0.20		
TFI T1			−0.07	−0.19	ΔTFI (T3-T1)			0.09	0.28
Interaction T1[a]	0.01	0.11	0.01	0.25	Interaction T1[b]	0.02**	0.51**	0.01*	0.37
General safety T1	0.04	0.06	0.04	0.06	General safety T1	−0.00	−0.01	0.02	0.03
SET T1	0.01	0.02			ΔSET (T3-T1)	0.00	0.01		
TFI T1			−0.02	−0.06	ΔTFI (T3-T1)			0.02	0.08
Interaction T1[a]	−0.01	−0.24	−0.00	−0.02	Interaction T1[b]	0.01[c]	0.29[c]	0.00	0.10
Unsafe location T1	0.27	0.28	0.28[c]	0.29	Unsafe location T1	0.31[c]	0.32[c]	0.31[c]	0.32
SET T1	0.07	0.12			ΔSET (T3-T1)	−0.03	−0.06		
TFI T1			0.06	0.15	ΔTFI (T3-T1)			−0.04	−0.10
Interaction T1[a]	0.00	0.05	0.00	0.06	Interaction T1[b]	−0.00	−0.00	0.00	0.04
Harassment T1	0.07	−0.07	−0.13	−0.13	Harassment T1	−0.08	−0.08	−0.08	−0.08
SET T1	0.05	0.08			ΔSET (T3-T1)	−0.03	−0.06		
TFI T1			0.06	0.14	ΔTFI (T3-T1)			−0.05	−0.12
Interaction T1b[a]	−0.01	−0.08	0.01	0.19	Interaction T1[b]	0.00	0.03	0.00	0.04
Additional support T1	0.18	0.21	0.18	0.21	Additional support T1	0.16	0.19	0.15	0.18
SET T1	0.02	0.12			ΔSET (T3-T1)	−0.01	−0.04		
TFI T1			0.03[c]	0.28	ΔTFI (T3-T1)			−0.01	−0.10
Interaction T1b[a]	−0.10	−0.22	−0.01	−0.22	Interaction T1[b]	0.01	0.10	0.00	0.03
Behavior incidents T1	0.41**	0.51**	0.51**	0.64	Behavior incidents T1	−0.45**	0.57**	0.41**	0.51
SET T1	−0.02[c]	−0.24[c]			ΔSET (T3-T1)	0.02*	0.27		
TFI T1			0.00	0.02	ΔTFI (T3-T1)			0.01	0.14
Interaction T1[c]	−0.02	−0.20	0.01[c]	0.27	Interaction T1[b]	0.01	0.14	−0.00	−0.06

Note: *$p < .05$; **$p < .01$.

Abbreviations: SET, School-wide Evaluation Tool; TFI, Tiered Fidelity Inventory.

[a]$0.05 < p < .1$.

[b]SET or TFI at T1 × student outcome variable T1.

[c]ΔSET or ΔTFI at T1 × student outcome variable.

students' social well-being at individual schools. This was also the case for the number of behavior incidents ($\beta = 0.51$). On the next two rows, we controlled for the level of fidelity. We saw no effect of the predictors SET or TFI at T1 on Well-being at T3, nor on any other variable. On the fourth row, the interaction effects are presented. The effects displayed in the third and fourth column, indicate the extent to which the stability depended on the level of the SET or TFI score. For none of the variables, the interaction effect was statistically significant.

We repeated these analyses using *changes* in fidelity scores instead of level of fidelity. Again, Well-being changed from T1 to T3 ($\beta = 0.49$; $p < 0.01$). In contrast to the result for the absolute level of fidelity, the effect of the interaction of change in fidelity and Well-being at T3 was significant ($\beta = 0.51$; $p < 0.01$). This indicates that the change in students' social well-being depended on the changes in fidelity. Behavior incidents showed a similar, though slightly different pattern: a significant change in the number of behavior incidents occurred, which was predicted by the change in fidelity ($\beta = 0.27$; $p < 0.05$). Other variables did not show significant changes.

For the TFI, results were similar. Well-being and Behavior incidents changed significantly. Other variables did not show significant change. TFI total scores were not related with student outcomes at T3. Changes in Well-being were significantly related to changes in TFI scores, indicating that students' social well-being increased at schools with increasing levels of implementation fidelity. In contrast to the SET, there was no significant relation between Behavior incidents and changes in TFI scores, indicating that the number of behavior incidents did not decrease at schools with increasing levels of fidelity.

10 | DISCUSSION

Little is known about the effects of implementing SWPBIS Tier 1 in the Netherlands. In this longitudinal study, we examined to what extent fidelity of SWPBIS implementation at Tier 1 in Dutch elementary schools was related to students' perceptions of social safety, the prevalence of behavior incidents, and the percentage of students receiving additional support for behavior, over 3 years. Our findings showed that fidelity scores, measured with both TFI and SET, and the percentage of students stating there were unsafe locations in and around school improved significantly from the first year to the third year. Students' well-being, general feelings of safety, harassment, behavior incidents, and students receiving additional support did not change significantly, though means scores of these variables headed in a similar, positive direction. We conducted two different analyses: the ANOVA repeated measures analyses (Table 2) were used to measure change over time in group means for fidelity and outcome variables. And second, we conducted regression analyses (Table 3) to detect patterns in individual school scores, controlling both for fidelity and change in fidelity. The ANOVA repeated measures showed a decrease in unsafe locations in and around schools, suggesting that students perceived their school as a more safe place to be. The regression analyses showed a decrease in behavior incidents, and an increase of student well-being. One could argue that these findings indicate that students increasingly perceived school as a safe place.

Although many studies (e.g., Simonsen et al., 2012) have shown that fidelity of implementation is crucial for achieving positive outcomes, in our study, a strong, unambiguous relation between fidelity and student outcomes was not found. Changes in fidelity were related to an increase in students' social well-being and a decrease in the number of behavior incidents, indicating that if a school strongly improved on fidelity of implementation, positive outcomes for social well-being and behavioral incidents were also likely to be seen. However, these results need to be carefully interpreted due to the absence of a control group. Other factors could also have influenced the changes found, such as maturation or staff turn-over.

Most likely, the composition of the sample influenced the results found. Most schools were already implementing SWPBIS, except nine schools that started 1 month before study onset. Nese et al. (2019) stated that the average period for elementary schools to reach adequate implementation was 2 years. In our sample, 36% of schools had implemented SWPBIS for more than 2 years at study onset. Bradshaw et al. (2009) saw organizational changes reaching significance at the end of Year 3 (21% of the schools in our sample at study onset). These findings

suggest that positive effects could already have been established in participating schools before the study started. This may be the reason why we did not see significant changes in most student outcomes, although we cannot be sure, as there were no pre-SWPBIS data (for both fidelity and student outcomes) available for these schools. For the nine schools that just had started implementing SWPBIS at study onset, we saw student outcomes improving over time, but the number of such starting schools was too small to draw solid conclusions from those data.

In our study, student outcome results were compared for a 3-year interval. Every year, student population changes due to students entering and leaving school. Thus the group of students in Year 1 (T1) was not the same as the group of students in Year 3 (T3). However, in our study, not the individual student, but the school was the unit of analysis. As data from large groups of students were aggregated at school level, these changes in student population are not likely to have significantly affected the outcomes of this study.

Reflecting on the outcomes of this study, another phenomenon that needs to be considered is a ceiling effect. Dutch schools perform relatively well. Most elementary students (94%–97% in 2010–2018) and teachers (94%–96%) feel safe at school, and not many major violent incidents occur (Nelen et al., 2018). In our study, at baseline (T1), students' social well-being was good and 86% of students generally felt safe at school. These figures seemed to leave not much room for improvement. However, creating safe schools remains important. For example, a 1% increase in social well-being or general feeling of safety would positively affect the life of more than 15,000 Dutch elementary students, which would make the effort of implementing SWPBIS worthwhile.

11 | LIMITATIONS AND FUTURE DIRECTIONS

For the first time, implementation of SWPBIS in Dutch schools was systematically examined in an exploratory study. Approximately 19% of all Dutch SWPBIS elementary schools participated. Several limitations should be noted. Despite multiple efforts, no control group could be assembled. Free lectures on data-based decision making were offered, but non-SWPBIS schools saw no gain in participating in a 3-year study. As a result, data from SWPBIS schools could not be compared with data from non-SWPBIS schools. Another research design often used, a pretest-posttest design, was also not an option, because all schools had already begun implementing SWPBIS at study onset. Instead, over a period of 3 years, we collected data on fidelity of implementation and student outcomes in 66 elementary schools that were implementing SWPBIS. For student outcomes, approximately 58% of partici-pating schools succeeded in providing the data requested. This presented us with a fait accompli of 42% missing data for student outcomes. For fidelity of implementation, the first author organized the data collection, and there were no data missing. Although we accounted for the missing data by examining if missing data were selective or not, and, based on this analysis, could conclude that the nonresponse was random and not selective, the large proportion of missing data negatively affected the power of our analyses. We carefully chose our outcome mea-sures, reminded schools via email to send in their data, and provided them with examples. Still, this type of research seems to be demanding for schools when data collection is not facilitated by researchers (Veerman et al., 2019).

Another limitation that needs to be addressed is the use of behavior incidents data in this study. First, although data-driven decision making is a distinct critical feature of SWPBIS, many Dutch SWPBIS schools struggle with collecting data on behavior incidents (Nelen et al., 2020b). Using prevalence or type of behavior incident data to develop preventive interventions was relatively new to them. Thirty percent of participating schools did not collect data on behavior incidents. Second, schools used different methods to collect data on behavior incidents. There-fore, we asked schools only to provide data on the number and location (in or outside class) of major and minor problem behaviors. In this study, we could not support schools in collecting behavior incidents data, other than providing them with examples each time we asked for their data. Third, reaching team agreement on what kind of incidents to log, and when and how to record them is hard to achieve in daily practice and often subject to fluctuations. Thus, it is unknown whether there was consistency in data collection within schools from 1 year to the next. Since this also applies for schools collecting data on ODRs, as reported in large effect studies such as

Bradshaw et al. (2010), we considered behavior incidents to be a similar–but not identical–outcome measure. And last, the findings for the number of behavior incidents could also have been influenced by the fact that we counted both major and minor incidents. According to Vincent et al. (2009), minor incidents often are not consistently reported. In U.S. research, usually only major incidents are taken into account. However, the analyses we performed were the best fit for the current situation in Dutch schools. As more and more Dutch schools start using the SWIS, this will ease analyzing behavioral data in future research.

In the Netherlands, several modalities exist for supporting schools in implementing SWPBIS. It is possible that student outcomes will vary depending on what kind of support a school receives. In our analyses, we could not use different forms of support as a covariate, because support was subject to many changes. Some schools received support from a SWPBIS coach all the time and others only at the beginning, schools changed coaches, or schools started with a SWPBIS coach and switched to network support. In future research this information should be collected and taken into account.

Despite these limitations, this study opened the way for further research and building of evidence regarding the use of SWPBIS in the Netherlands. According to Horner et al. (2010), documenting the evidence base for SWPBIS is complex, as it is a "large constellation of systems and practices" (p. 5). Implementing SWPBIS is considered to be a school development process, with many factors influencing the outcomes. If SWPBIS alters school organizations (Bradshaw et al., 2009), it is most likely that it also affects the faith a school community has in their ability to change for the better. Most likely, implementing SWPBIS establishes a kind of "school efficacy," and this growing faith also contributes to achieving positive outcomes.

12 | CONCLUSIONS AND IMPLICATIONS FOR PRACTICE

In an exploratory longitudinal study on effects of SWPBIS implementation in 66 Dutch elementary schools, a significant increase in fidelity scores, and a significant decline in the percentage of students stating there were unsafe locations in and around school were found. Changes in fidelity were related to an increase in students' social well-being and a decrease in the number of behavior incidents, indicating that if a school strongly improved on fidelity of implementation, positive outcomes for social well-being and behavior incidents were also likely to be seen. These results imply that taking care to implement SWPBIS with fidelity is important. For schools starting to implement SWPBIS, working on fidelity can make a change in achieving positive student outcomes. For schools that have already reached an acceptable level of implementation, sustaining fidelity can contribute to the continuation of positive student outcomes. In our study, schools highly valued the yearly measurement of fidelity. Fidelity instruments serve multiple purposes (McIntosh et al., 2017). Together with a SWPBIS coach (e.g., a school psychologist), school leadership teams can obtain a clear overview of what core features of SWPBIS they have already realized, and what is yet to be done toward full implementation. As TFI measurements come with an action planning tool, schools can learn to use their data to identify needs for improvement in their systems. Measuring and analyzing student outcomes systematically can provide schools with information on their output. Fidelity therefore has important links with practice on many levels.

CONFLICT OF INTERESTS
The authors declare that there are no conflict of interests.

REFERENCES

Bradshaw, C. P., Koth, C. W., Bevans, K. B., Lalongo, N., & Leaf, P. J. (2008). The Impact of schoolwide Positive Behavioral Interventions and Supports (PBIS) on the organizational health of elementary schools. *School Psychology Quarterly*, *23*(4), 462–473. https://doi.org/10.1037/a0012883

Bradshaw, C. P., Koth, C. W., Thornton, L. A., & Leaf, P. J. (2009). Altering school climate through schoolwide Positive Behavioral Interventions and Supports: Findings from a group-randomized effectiveness trial. *Prevention Science*, *10*(2), 100–115. https://doi.org/10.1007/s11121-008-0114-9

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide Positive Behavioral Interventions and Supports on student outcomes. Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions*, *12*(3), 133–148. https://doi.org/10.1177/1098300709359084

Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of schoolwide Positive Behavioral Interventions and Supports on child behavior problems. *Pediatrics*, *130*(5), E1136–E1145. https://doi.org/10.1542/peds.2012-0243

Center on Positive Behavioral Interventions and Supports. (2015). Positive Behavioral Interventions and Supports (PBIS) Implementation Blueprint. University of Oregon. Retrieved from: www.pbis.org

Cohen, J., McCabe, E. M., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *Teachers College Record*, *111*(1), 180–213.

Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice*, *19*(5), 531–540.

Greenwood, C. R., Kratochwill, T. R., & Clements, M. (2008). Schoolwide prevention models, *Lessons learned in elementary schools*. The Guilford Press.

Horner, R. H., Sugai, G., & Anderson, C. M. (2010). Examining the evidence base for School-Wide Positive Behavior Support. *Focus on Exceptional Children*, *42*(8), 1–14. https://doi.org/10.17161/fec.v42i8.6906

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing School-Wide Positive Behavior Support in elementary schools. *Journal of Positive Behavior Interventions*, *11*(3), 133–144. https://doi.org/10.1177/1098300709332067

Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. B. (2004). The Schoolwide Evaluation Tool (SET): A research instrument for assessing School-Wide Positive Behavior Support. *Journal of Positive Behavior Interventions*, *6*(3), 3–12. https://doi.org/10.1177/10983007040060010201

Kutsyuruba, B., Klinger, D. A., & Hussain, A. (2015). Relationships among school climate, school safety, and student achievement and well-being: A review of the literature. *Review of Education*, *3*(2), 103–135. https://doi.org/10.1002/rev3.3043

Lewis, T. J., Barrett, S., Sugai, G., Horner, R. H., Mitchell, B. S., & Starkey, D. (2016). *Training and professional development blueprint for positive behavioral interventions and supports*. National Technical Assistance Center on Positive Behavior Interventions and Support. Retrieved on January 2021 from https://www.pbis.org

Marlet, G., Poort, J., & van Woerkens, C. (2009). De baat op straat. Het effect van de investeringen van woningcorporaties op overlast, onveiligheid en verloedering in de buurt [Street benefits. The effect of housing corporations' investements on nuisance, unsafety, and degeneration in the neighborhood]. Atlas voor gemeenten.

May, S., Ard, W., Todd, A., Horner, R., Glasgow, A., Sugai, G., & Sprague, J. (2010). *School-Wide Information System v4.5.217*, Eugene, Oregon: University of Oregon.

McIntosh, K., Bennett, J. L., & Price, K. (2011). Evaluation of social and academic effects of School-Wide Positive Behaviour Support in a Canadian school district. *Exceptionality Education International*, *21*(1), 46–60.

McIntosh, K., Ellwood, K., McCall, L., & Girvan, E. J. (2018). Using discipline data to enhance equity in school discipline. *Intervention in School and Clinic*, *53*(3), 146–152. https://doi.org/10.1177/1053451217702130

McIntosh, K., Filter, K. J., Bennett, J. L., Ryan, C., & Sugai, G. (2010). Principles of sustainable prevention: Designing scale-up of School-Wide Positive Behavior Support to promote durable systems. *Psychology in the Schools*, *47*(1), 5–21. https://doi.org/10.1002/pits.20448

McIntosh, K., Massar, M. M., Algozzine, R. F., Peshak George, H., Horner, R. H., Lewis, J. L., & Swain-Bradway, J. (2017). Technical adequacy of the SWPBIS Tiered Fidelity Inventory. *Journal of Positive Behavior Interventions*, *19*(1), 3–13. https://doi.org/10.1177/1098300716637193

McIntosh, K., Mercer, S. H., Hume, A. E., Frank, J. L., & Turri, M. G. (2013). Factors related to sustained implementation of School-Wide Positive Behavior Support. *Exceptional Children*, *79*(3), 293–311. https://doi.org/10.1177/1098300715599737

McIntosh, K., Mercer, S. H., Nese, R. N. T., Strickland-Cohen, M. K., & Hoselton, R. (2015). Predictors of sustained implementation of School-Wide Positive Behavioral Interventions and Supports. *Journal of Positive Behavior Interventions*, *18*(4), 209–218. https://doi.org/10.1177/1098300715599737

Mercer, S. H., McIntosh, K., & Hoselton, R. (2017). Comparability of fidelity measures for assessing Tier 1 School-Wide Positive Behavioral Interventions and Supports. *Journal of Positive Behavior Interventions*, *19*(4), 195–204. https://doi.org/10.1177/1098300717693384

Mooij, T., De Wit, W., & Fettelaar, D. (2011). *Sociale veiligheid in en rond scholen. Primair en Voortgezet Onderwijs 2006 - 2010 [Social safety in and around schools for Primary and secondary education 2006 - 2010]*, Radboud Universiteit, ITS.

Nelen, M. J. M., Blonk, A., Scholte, R. H. J., & Denessen, E. (2020a). School-Wide Positive Behavior Interventions and Supports: fidelity of Tier 1 implementation in 117 Dutch schools. *Journal of Positive Behavior Interventions*, 22(3), 156–166. https://doi.org/10.1177/10983007198796

Nelen, M. J. M., Willemse, T. M., van Oudheusden, M. A., & Goei, S. L. (2020b). Cultural challenges in adapting SWPBIS to a Dutch context. *Journal of Positive Behavior Interventions*, 22(2), 105–115. https://doi.org/10.1177/1098300719876096

Nelen, W. B. L., De Wit, W., Golbach, M., van Druten, L., Deen, C., & Scholte, R. (2018). *Sociale veiligheid in en rond scholen [Social safety in and around schools]*. Nijmegen, The Netherlands: Praktikon.

Nese, R. N. T., Nese, J. F. T., McIntosh, K., Mercer, S. H., & Kittelman, A. (2019). Predicting latency of reaching adequate implementation of Tier I School-Wide Positive Behavioral Interventions and Supports. *Journal of Positive Behavior Interventions*, 21(2), 106–116.

Nijs, M. M., Bun, C. J. E., Tempelaar, W. M., de Wit, N. J., Burger, H., Plevier, C. M., & Boks, M. P. M. (2014). Perceived school safety is strongly associated with adolescent mental health problems. *Community Mental Health Journal*, 50(2), 127–134. https://doi.org/10.1007/s10597-013-9599-1

Office of Special Education Programs (OSEP) Technical Assistance Center on Positive Behavioral Interventions and Supports (2015). Positive Behavioral Interventions and Supports (PBIS) Implementation Blueprint: Part 1 – Foundations and Supporting Information. University of Oregon. Retrieved from www.pbis.org

Organisation for Economic Co-operation and Development (OECD). (2011). *Education at a glance 2011*, OECD indicators, OECD Publishing.

Sailor, W., Dunlap, G., Sugai, G., & Horner, R. H. (2009). *Handbook of Positive Behavior Support*. Springer.

Schulte, A. C., Easton, J. E., & Parker, J. (2009). Advances in treatment integrity research: Multidisciplinary perspectives on the conceptualization, measurement, and enhancement of treatment integrity. *School Psychology Review*, 38, 460–475.

Simonsen, B., Eber, L., Black, A. C., Sugai, G., Lewandowski, H., Sims, B., & Myers, D. (2012). Illinois statewide Positive Behavioral Interventions and Supports: Evolution and impact on student outcomes across years. *ournal of Positive Behavior Interventions*, 14(1), 5–16. https://doi.org/10.1177/1098300711412601

Smith, T. B., Domenech Rodríguez, M., & Bernal, G. (2011). Culture. *Journal of Clinical Psychology*, 67(2), 166–175. https://doi.org/10.1002/jclp.20757

Sugai, G., Horner, R. H., & McIntosh, K. (2008). Best practices in developing a broad-scale system of support for schoolwide positive behavior support. In (Eds.) Thomas, A. & Grimes, J. P., *Best Practices in School Psychology V* (pp. 765–780). National Association of School Psychologists.

Sugai, G., O'Keeffe, B. V., & Fallon, L. M. (2012). A contextual consideration of culture and School-Wide Positive Behavior Support. *Journal of Positive Behavior Interventions*, 14(4), 197–208. https://doi.org/10.1177/1098300711426334

Sugai, G., & Simonsen, B. (2012). Positive Behavioral Interventions and Supports: history, defining features, and misconceptions. Retrieved from: www.pbis.org/resource/positive-behavioral-interventions-and-supports-history-defining-features-and-misconceptions. [accessed 19 January 2021].

Sørlie, M., & Ogden, T. (2015). School-Wide Positive Behavior Support–Norway: impacts on problem behavior and classroom climate, *International Journal of School & Educational Psychology*, pp. 1–16. https://doi.org/10.1080/21683603.2015.1060912

Veerman, J. W., Hendriks, L., van Huijgevoort, H., Blonk, A., & Dollevoet, T. (2019). *Met andere ogen. Denken en doen in praktijkgestuurd onderzoek [Looking differently. Thinking and acting in practice-based research]*. Oud-Turnhout/'s Hertogenbosch: Gompel & Svacina.

Vincent, C., Horner, R., & May, S. (2009). *What are the patterns of office discipline referrals across grade levels?* Evaluation brief retrieved from https://www.pbis.org/resource/what-are-the-patterns-of-office-discipline-referrals-across-grade-levels. [accessed 19 January 2021].

Waasdorp, T. E., Bradshaw, C. P., & Leaf, P. J. (2012). The impact of Schoolwide Positive Behavioral Interventions and Supports on bullying and peer rejection: A Randomized Controlled Effectiveness Trial. *Archives of Pediatrics & Adolescent Medicine*, 166(2), 149–156. https://doi.org/10.1001/archpediatrics.2011.755

Yeung, A. S., Craven, R. G., Mooney, M., Tracey, D., Barker, K., Power, A., Dobia, B., Chen, Z., Schofield, J., Whitefield, P., & Lewis, T. J. (2016). Positive behavior interventions: The issue of sustainability of positive effects. *Educational Psychology Review*, 28(1), 145–170. https://doi.org/10.1007/s10648-015-9305-7