GION
ONDERWIJS/ONDERZOEK

Effective classroom management strategies and classroom management programs for educational practice

A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioural, emotional, and motivational outcomes

Hanke Korpershoek
Truus Harms
Hester de Boer
Mechteld van Kuijk
Simone Doolaard

Effective classroom management strategies and classroom management programs for educational practice

Effective classroom management strategies and classroom management programs for educational practice

A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioural, emotional, and motivational outcomes.

Hanke Korpershoek, Truus Harms, Hester de Boer, Mechteld van Kuijk, Simone Doolaard

Effective classroom management strategies and classroom management programs for educational practice

A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioural, emotional, and motivational outcomes.

Table of contents

Nederlandstalige samenvatting

In dit rapport wordt verslag gedaan van een meta-analyse van de effectiviteit van verschillende klassenmanagementstrategieën en klassenmanagementprogramma's in het primair onderwijs. Het onderzoek is gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Centraal stond de vraag in hoeverre bepaalde klassenmanagementstrategieën en klassenmanagementprogramma's de leerprestaties van leerlingen, het leerlinggedrag, de sociaal-emotionele ontwikkeling van leerlingen en de motivatie van leerlingen verbeteren.

De definitie van klassenmanagement die in het onderzoek gehanteerd wordt is gebaseerd op het werk van Evertson en Weinstein (2006). Zij definiëren klassenmanagement als "de handelingen die leerkrachten ondernemen om een omgeving te creëren die academisch en sociaal-emotioneel leren ondersteunt en faciliteert" (pp. 4-5). Deze definitie benadrukt de verantwoordelijkheid van de leerkracht en legt bovendien een relatie tussen klassenmanagement en verschillende leeruitkomsten van leerlingen. Goed klassenmanagement betekent dus de leerlingen de ruimte geven om te leren. Leren op het cognitieve vlak, wat zich uit in de prestaties op bijvoorbeeld taal, lezen, rekenen en wereldoriëntatie, maar ook leren op het niet-cognitieve vlak, zoals het leren omgaan met andere leerlingen.

In de meta-analyse zijn uiteindelijk 47 geschikte studies geselecteerd waarin in totaal 54 klassenmanagementinterventies beschreven staan. Deze interventies voldoen in ieder geval aan de volgende eisen: ze zijn gedegen onderzocht en beschreven (o.a. gepubliceerd in peer-reviewed wetenschappelijke tijdschriften en gebruik makend van een experimentele of quasi-experimentele onderzoeksopzet), ze zijn gericht op de hele groep leerlingen (en niet op één of enkele leerlingen), ze hebben uitkomstmaten op leerlingniveau gemeten (bv. leerlinggedrag), de interventies zijn uitgevoerd door de groepsleerkracht zelf (dus niet door schoolbegeleiders of onderzoekers) en de studies zijn verschenen tussen 2003 en 2013.

In de literatuur is een breed scala aan klassenmanagementstrategieën en klassenmanagementprogramma's gevonden. In sommige studies heeft de leerkracht een klassenmanagementstrategie uitgeprobeerd, bijvoorbeeld het belonen van goed gedrag. Daarnaast zijn schoolbrede klassenmanagementprogramma's (of interventies waarin klassenmanagement onderdeel is van het programma) meegenomen in de meta-analyse. Dit zijn programma's waarbij bijvoorbeeld op de hele school dezelfde gedragsregels afgesproken worden en waar extra lessen gegeven worden in 'hoe je met elkaar omgaat' (m.a.w. aandacht voor de sociaal-emotionele ontwikkeling van kinderen). De interventies zijn vervolgens

geclassificeerd op basis van vier categorieën. De vier categorieën verwijzen naar de focus van de interventies: gericht op (1) leerkrachtgedrag, (2) leerkracht-leerling relaties, (3) leerlinggedrag, en (4) de sociaal-emotionele ontwikkeling van leerlingen. Uiteraard vallen sommige interventies onder meerdere categorieën en sluiten deze categorieën elkaar niet uit. De effecten van de interventies in deze (of combinaties van deze) categorieën zijn onderzocht. Daarnaast is het effect onderzocht van vijf programma's die in meerdere studies (tenminste 3) aan de orde kwamen: Promoting Alternative Thinking Strategies (PATHS), Good Behavior Game (GBG), Zippy's Friends, School-Wide Positive Behavior Support (SWPBS) en Second Step. Van de eerste vier programma's zijn ook Nederlandse varianten op de markt gebracht.

Over het algemeen bleken klassenmanagementinterventies een positief effect te hebben op de leerlingen. Dat wil zeggen, het gemiddelde effect van alle interventies op de leerlinguitkomsten was positief en significant (de gemiddelde effectgrootte Hedges's g was 0,22). De effecten van de interventies op de verschillende uitkomstmaten afzonderlijk (leerprestaties, leerlinggedrag, sociaal-emotionele ontwikkeling van leerlingen, motivatie, en een categorie overige relevante uitkomstmaten zoals time-on-task) verschillen nauwelijks van elkaar; behalve voor motivatie, waarvoor geen significant effect gevonden is. De interventies waarbij de focus (onder meer) lag op het verbeteren van de sociaal-emotionele ontwikkeling van leerlingen waren iets effectiever dan interventies zonder deze focus, en dat betrof met name de sociaal-emotionele uitkomstmaten. Wat betreft de leerprestaties van de leerlingen bleken de interventies waarbij de focus (onder meer) lag op het verbeteren van leerkrachtgedrag (bv. door bepaalde klassenmanagementvaardigheden aan te leren) iets effectiever.

Uit de analyse van de vijf programma's die in meerdere studies beschreven stonden bleek dat deze programma's over het algemeen even effectief waren (kleine tot middelgrote effecten), zij het dat deze de effecten wat verschilden voor de afzonderlijke uitkomstmaten. Het SWPBS-programma vormde hierop een uitzondering: voor dit programma zijn nauwelijks significante effecten op leerlinguitkomsten gevonden. Het PATHS-programma (in Nederland bekend onder de naam *Programma Alternatieve Denkstrategieën, PAD*) viel verder op doordat het een groter effect liet zien dan de andere programma's op de sociaal-emotionele ontwikkeling van leerlingen.

Prologue

Teachers play a fundamental role in the cognitive and social-emotional development of children by giving them the opportunity to learn. Effective classroom management sets the stage for this learning. Without it, classrooms are disorganized and chaotic, and very little academic learning can happen (Elias & Schwab, 2006, p. 309).

The findings of numerous studies have shown that teachers play a key role in shaping effective education. Effective classroom management is a requirement for effective education. In this study, we provide an overview of classroom management strategies and classroom management programs for (new) teachers in primary education to help them develop ways to effectively manage their classrooms and to identify interventions that have the potential to prevent classroom management difficulties.

1   Introduction

1.1 Introduction

Effective education refers to the degree to which schools are successful in accomplishing their educational objectives. The findings of numerous studies have shown that teachers play a key role in shaping effective education (Hattie, 2009). The differences in achievement between students who spend a year in a class with a highly effective teacher as opposed to a highly ineffective teacher are startling. Marzano (2003) synthesized 35 years of research on effective schools and found the following results. Consider the following case: a student attends an average school and has an average teacher for two years. At the end of these two years, the student's achievement will be at the 50th percentile. If the same student attends an ineffective school and has an ineffective teacher, the student's achievement will have drop to the 3rd percentile after two years. If the student attends an effective school but has an ineffective teacher, his or her achievement after two years will have dropped to the 37th percentile. An individual teacher can produce powerful gains in student learning.

Effective teaching and learning cannot take place in poorly managed classrooms (Jones & Jones, 2012; Marzano, Marzano, & Pickering, 2003; Van de Grift, Van der Wal, & Torenbeek, 2011). Effective classroom management strategies (hereafter abbreviated to CMS) support and facilitate effective teaching and learning. Effective classroom management is generally based on the principle of establishing a positive classroom environment encompassing effective teacher-student relationships (Wubbels, Brekelmans, Van Tartwijk, & Admiraal, 1999). Evertson and Weinstein (2006) define *classroom management* as "the actions teachers take to create an environment that supports and facilitates both academic and social-emotional learning" (pp. 4-5). This definition concentrates on the responsibility of the teacher and relates the use of classroom management strategies to multiple learning goals for students.

Following this definition, effective CMS seem to focus on preventive rather than reactive classroom management procedures (Lewis & Sugai, 1999). An example of a widely used – and generally effective – preventive strategy among teachers in primary education is that classroom rules are negotiated instead of imposed (Marzano et al., 2003). Teachers, however, also frequently use reactive strategies (e.g., punishing disruptive students; Rydell & Henricsson, 2004; Shook, 2012), whereas it is unclear whether these strategies effectively change student behaviour. This may be caused by a lack of knowledge about the effectiveness of preventive strategies (e.g., Peters, 2012), or by a lack of belief in their effectiveness. Teachers do not always believe in the effectiveness of particular strategies despite ample empirical evidence that the strategy has been implemented successfully in many classrooms (e.g., Smart & Brent, 2010). One example is that beginning teachers are generally advised to

be as strict as possible in the first week of their internship and then slowly to become less authoritarian, whereas first establishing positive teacher-student relationships has been proven far more effective in regulating student behaviour (e.g., Bohn, Roehrig, & Pressley, 2004). O'Neill and Stephenson (2012) emphasize that completing focused coursework units on classroom management in teacher training programs leads to increased feelings of preparedness, familiarity, and confidence in using CMS among student teachers. However, they stress that student teachers reported that they were confident in using only half of the strategies they were familiar with, and that they only felt partially prepared to manage the misbehaviour of students. When teachers feel uncertain about using preventive strategies, for instance, negotiating about classroom rules, they often keep using the (presumably less effective) reactive strategies (Rydell & Henricsson, 2004; Woodcock & Reupert, 2012).

Evidently, mastering effective CMS is a basic competence for all teachers ("Wet op de beroepen in het onderwijs", freely translated as "Professions in Education Act"). Klamer-Hoogma (2012) stresses that good teachers need to master a broad range of CMS, and that teacher training programs should provide student teachers with a large "toolbox" of CMS from which they can pick and apply particular strategies when necessary. Which strategies should (at least) be part of this so-called toolbox in *current* educational settings is still unclear. The reason for this is that the books that are used in teacher training programs (e.g., Jones & Jones, 2012; Klamer-Hoogma, 2012; Teitler & Van Brussel, 2012) generally refer to studies that were conducted decades ago or used anecdotal evidence rather than empirical evidence. However, daily practice in education has changed rapidly. It is increasingly characterized by student-centred approaches to learning (as opposed to teacher-centred), with a large emphasis on students' metacognitive skills (e.g., self-regulated learning strategies; Dignath, Büttner, & Langfeldt, 2008) and cooperative learning (e.g., Kagan, 2005; Wubbels, Den Brok, Veldman, & Van Tartwijk, 2006). Moreover, more and more technology is finding its way into classrooms, for example, the use of interactive whiteboards, tablets, and laptops (Schussler, Poole, Whitlock, & Evertson, 2007). These changes presumably have had a large impact on the demands placed on teachers' classroom management skills (e.g., rules and procedures to facilitate cooperative learning). Although, to the best of our knowledge, no studies have been conducted to explicitly compare the effectiveness of particular CMS in more traditional versus more modern classrooms, an up-to-date overview of studies conducted in the last decade is expected to provide insight into which CMS have been proven (still) to be effective in modern classrooms.

1.2 Objective and research question

Our main objective was to conduct a meta-analysis of the effects of various CMS and classroom management programs (hereafter called "CMP") aimed at improving students' behaviour and enhancing their academic performance in primary education. In line with

Evertson and Weinstein's (2006) definition of classroom management, we focused on the literature on CMS/CMP that support and facilitate both academic and social-emotional learning. As a result, the meta-analysis included studies conducted to examine the effects of CMS/CMP on various student outcomes, namely, academic outcomes (e.g., student performance), behavioural outcomes, social-emotional outcomes, and motivational outcomes. The following research question was formulated:

Which classroom management strategies and classroom management programs effectively support and facilitate academic, behavioural, social-emotional, and/or motivational outcomes in primary education?

This question was addressed by performing a systematic review of the peer-reviewed classroom management literature published between 2003 and 2013. The results of the meta-analysis give an overview of contemporary effective CMS/CMP that improve student outcomes. This knowledge base supports (beginning) teachers in effectively managing their classrooms in current educational settings. The study was concentrated on whole-classroom interventions, implemented by teachers in their own classrooms (including school-wide interventions).

2   Theoretical framework

In this chapter, we present the definitions of classroom management that formed the basis for our study. We then discuss different types of CMS; various criteria are used in the field of classroom management to distinguish between different CMS. Next, we provide a summary of the results of previously conducted review studies and meta-analyses of the effects of these CMS/CMP on different student outcomes. After this, five frequently implemented CMP are discussed to provide the reader with examples of current programs.

2.1 Definition of classroom management

Evertson and Weinstein (2006) refer in their definition of classroom management to the actions teachers take to create a supportive environment for the academic and social-emotional learning of students. They describe five types of actions. In order to attain a high quality of classroom management, teachers must (1) develop caring, supportive relationships with and among students and (2) organize and implement instruction in ways that optimize students' access to learning. The importance of developing favourable teacher-student relationships is also expressed by Marzano et al. (2003). Additionally, Evertson and Weinstein (2006) state that teachers should (3) encourage students' engagement in academic tasks, which can be done by using group management methods (e.g., by establishing rules and classroom procedures, see Marzano et al., 2003). Teachers must (4) promote the development of students' social skills and self-regulation. Marzano et al. (2003) refer to this as making students responsible for their behaviour. Finally, Evertson and Weinstein (2006) state that teachers should be able to (5) use appropriate interventions to assist students with behaviour problems. The last two actions proposed by Evertson and Weinstein (2006) indicate that effective classroom management improves student behaviour. Hence, classroom management is an ongoing interaction between teachers and their students. Brophy (2006) presents a similar definition: "*Classroom management refers to actions taken to create and maintain a learning environment conducive to successful instruction (arranging the physical environment, establishing rules and procedures, maintaining students' attention to lessons and engagement in activities)*" (p. 17). Both definitions emphasize the importance of actions taken by the teacher to facilitate learning among the students.

2.2 Classroom management strategies and different classifications of CMS

As stated above, classroom management is about creating inviting and appealing environments for student learning. Classroom management strategies are tools that the

teachers can use to help create such an environment, ranging from activities to improve teacher-student relationships to rules to regulate student behaviour. Only when the efforts of management fail should teachers have to resort to reactive, controlling strategies. Therefore, it is important to distinguish between preventive and reactive classroom management strategies. That is, there is a difference between strategies used to prevent behaviour problems and strategies used to respond to problem behaviour (see also Lane, Menzies, Bruhn, & Cmobori, 2011, or the Dutch translation of this classroom management book by Lane & Kuiper, 2012). For example, the establishment of rules and procedures and favourable teacher-student relationships are considered preventive strategies, whereas disciplinary interventions such as giving warnings or punishments are considered reactive strategies. Although it is generally assumed that preventive strategies are more effective than reactive strategies, reactive strategies are sometimes needed to reduce disruptive or other undesired student behaviour when preventive strategies do not work (Marzano et al., 2003). In a similar vein, Froyen and Iverson (1999) used the concepts management of content (e.g., space, materials, equipment, movement, and lessons) and management of covenant (e.g., social dynamics and interpersonal relationships) for preventive strategies, and management of conduct (e.g., disciplinary problems) for reactive strategies when referring to classroom management.

A separate group of CMS are group contingencies, which represent various reinforcement strategies aimed at improving student behaviour or performance. These include preventive and reactive strategies. These group contingencies can be classified into three types (as discussed in Kelshaw-Levering, Sterling-Turner, & Henry, 2000): independent, interdependent, and dependent group contingencies. Independent group contingencies refer to reinforcement interventions that apply the same assessment criteria and reinforcements to each child (e.g., all children should pass the same swimming test before they get a diploma). Dependent group contingencies, on the other hand, refer to interventions that require a single student (or a few students) to reach a designated criterion in order for the whole group to receive reinforcement (e.g., when a student attains a 100 percent score on a test, the teacher will hand out sweets to the entire class). Interdependent group contingencies require the whole student group to reach a designated criterion in order to receive reinforcement (e.g., group members need to collaborate on a team project and the entire team receives a grade for their end product).

Returning to the preventive-reactive classification, both preventive strategies and reactive strategies can be applied to the entire classroom population (e.g., by discussing classroom rules or giving group detention) or to individual students (e.g., by letting an easily distracted student sit alone during independent seatwork or placing a student temporarily outside the classroom when showing disruptive behaviour). We limited the current investigation to whole-class classroom management interventions, because the methods used to investigate strategies to improve individual students' behaviour (e.g., students with behavioural and/or emotional disorders) or to discipline individual students (e.g., move seat, isolation time out, detention) are usually single case studies – mostly with multiple baseline designs – which

cannot be combined with control group designs in the same meta-analysis. Without a control group, maturation effects cannot be detected. Particularly for social-emotional and behavioural outcomes, maturation effects are part of students' natural development (e.g., Erikson, 1968). Moreover, it seems that effective management of the whole classroom population (including adequate response to disruptive individual students) is a prerequisite for dealing with students requiring additional behavioural support (see Swinson, Woof, & Melling, 2003).

When comparing the above-mentioned classifications of classroom management strategies (preventive/reactive; management of content/convenant/conduct), we did not find a systematic classification of classroom management interventions that covers the whole range of classroom management dimensions following from Evertson and Weinstein's (2006) definition of classroom management (we consider their work to provide the most exhaustive description of what classroom management entails). Improving student behaviour (e.g., self-control) is an important goal in many classroom management programs nowadays, while this student component is underrepresented in the different classifications mentioned above. Moreover, in many interventions, both preventive and reactive strategies are used. Therefore, we propose the following classification ("types") of classroom management interventions, based on their primary focus:

1) Teachers' behaviour-focused interventions. The focus of the intervention is on improving teachers' classroom management (e.g., keeping order, introducing rules and procedures, disciplinary interventions) and thus on changing the teachers' behaviour. This type is a representation of the group management methods referred to by Evertson and Weinstein (2006). Both preventive and reactive interventions are included in this category.

2) Teacher-student relationship-focused interventions. The focus of the intervention is on improving the interaction between teachers and students (teacher-student interactions), thus on developing caring, supportive relationships. Only preventive interventions are included in this category. This type is a representation of the supportive teacher-student relation referred to by Evertson and Weinstein (2006). Interventions focusing on relations between students only (and not the relation between the teacher and the students) are not included here; these are classified as type 4 (see below).

3) Students' behaviour-focused interventions. The focus of the intervention is on improving student behaviour, for example, via group contingencies or by improving self-control among all students. Both preventive and reactive interventions are included in this category. This type is a representation of the students' self-regulation referred to by Evertson and Weinstein (2006), as well as Marzano et al. (2003), who refer to students' responsibility for their own behaviour.

4) Students' social-emotional development-focused interventions. The focus of the intervention is on improving students' social-emotional development, such as enhancing their feelings of empathy for other children. Both preventive and reactive interventions are included in this category. This type is a representation of the students' social skills referred to by Evertson and Weinstein (2006).

Evidently, some classroom management programs may fit into more than one of these categories; the types are not considered to be mutually exclusive. The proposed classification was used in the meta-analysis to identify the differential effects of different types of interventions. One particular type of intervention might be more effective than other types. Moreover, it is possible that broader interventions which have multiple foci may establish stronger effects than interventions that have one primary focus, or that a particular combination of foci may be more effective than other combinations.

2.3 Prior meta-analyses

Three relevant prior meta-analyses are summarized in this section. The study by Marzano et al. (2003) is the most recent meta-analysis of effective classroom management; it not only used the number of disruptions in classrooms as the outcome variable, but also included student engagement and student achievement as outcome measures. This meta-analysis has been widely cited in the classroom management literature used in teacher training programs. The following two (recent) meta-analyses are discussed here as they touch on different dimensions of classroom management. The study by Oliver, Wehby, and Reschly (2011) reports on the effects of classroom management practices on problem-student behaviour, and thus has a much narrower scope than Marzano et al. (2003) and the present meta-analysis. The meta-analysis by Durlak, Weissberg, Dymnicki, Taylor, and Schellinger (2011) concentrates on the effectiveness of social and emotional learning programs on various student outcomes. Such programs generally include classroom management components, which is why we include an overview of their findings here.

2.3.1 Marzano, Marzano, & Pickering (2003)
Marzano et al.'s (2003) study was based on 101 studies into effective classroom management, published between 1967 and 1997. The participants were primary and secondary school students; students in regular as well as special education were included. About half of the studies were based on a single subject, the other half on groups of students. The majority of the studies included in the meta-analysis focused on only one of the components of classroom management described below.

The findings revealed that the average number of classroom interruptions was evidently lower in classes where CMS were used effectively than in classrooms where these strategies were not used effectively. Marzano et al. found four general components of teachers' effective classroom management, most of which are included in Evertson and Weinstein's (2006)

definition of classroom management (see section 3.1). Marzano et al.'s four components are: (1) rules and procedures, (2) disciplinary interventions, (3) teacher-student relationships, and (4) mental set. The authors state that, in effectively managed classrooms, there are clear rules and procedures that express the expected behaviour. Rules refer to general behaviours relevant to how to treat each other (e.g., "We are kind to each other"); procedures refer to behaviours in specific situations, for example, the beginning of the school day (e.g., "We place the lunch box in the kitchen"), or transitions (e.g., "When you finish your task, you put it on the teachers' desk"). Disciplinary interventions are about how to deal with disruptive behaviour. These include strategies for punishment of negative behaviour (e.g., temporarily placing a student outside the classroom) and reward of positive behaviour (e.g., free time for playing games). The third component of effective classroom management includes techniques for teachers to establish appropriate levels of dominance and cooperation in the classroom in order to optimize teacher-student relationships. Examples are setting clear goals, showing interest in students' concerns, and interacting in an equitable and positive way. 'Mental set' includes two aspects: 'withitness' and 'emotional objectivity'. It includes the disposition of the teacher to quickly and accurately identify potential problem behaviour and to act on it immediately ('with it') in an emotionally objective way (that is, not getting angry or frustrated). In addition to these four components of teachers' classroom management, Marzano et al. emphasize the importance of making students responsible for their behaviour. Teachers can teach students strategies to self-monitor and control their behaviour; in other words, they can delegate authority to the students rather than guide their behaviour directly.

The researchers reported an effect size of Cohen's d = -0.76, 95% CI [-0.60; -0.93] for rules and procedures (based on ten studies). The interpretation of this effect size is as follows: in classrooms focused on effective use of rules and procedures, the average number of classroom interruptions was 0.76 standard deviations less than in classrooms that were not focused on these techniques. For disciplinary interventions, the effect size was d = -0.91 (CI not reported); for teacher-student relationships, d = -0.87, 95% CI [-0.74; -1.00]; for mental set, d = -1.29, 95% CI [-1.10; -1.49]; and for student responsibility, d = -0.69, 95% CI [-0.56; -0.83], based on sixty-eight, four, five, and twenty-eight studies, respectively. These are all high effect sizes. Effect sizes for each component of CMS were also presented for primary and secondary schools separately. The effect sizes and their subsequent confidence intervals indicated that a difference between these two school levels could only be found for disciplinary interventions, with a higher average effect size for primary school. Furthermore, Marzano et al. showed that observations of students' engagement (observer ratings) and measures of students' achievement levels were clearly higher in classes in which effective CMS were used than in classes in which effective management strategies were not used. The meta-analysis included seven studies in which the effects of CMS on engagement were measured and five studies in which the effects on achievement were measured; the results revealed average effects of 0.62 and 0.52 standard deviations higher, respectively. All effect

sizes reported above were significant ($p < 0.05$). Based on these results, Marzano et al. (2003) emphasized that classroom management is one of the "critical ingredients of effective teaching" (p. 6).

A limitation of Marzano et al.'s meta-analyses is that the authors did not report how they performed the literature search (i.e., what search terms and eligibility criteria were used) and how the meta-analysis was executed. As a result, the exact methods used to come to their findings remain obscure. For example, it is unclear how the authors came to the selected studies, whether the primary studies were experiments in which the effects of CMS were examined rather than correlational studies, and whether a control group was always used. Nonetheless, Marzano et al.'s results do suggest that CMS are an important instrument for creating an orderly and harmonious learning environment.

2.3.2 Oliver, Wehby, & Reschly (2011)

Oliver et al.'s (2011) review focuses on universal, whole-class classroom management procedures. They define whole-class procedures as "a collection of non-instructional classroom procedures implemented by teachers in classroom settings with all students for the purposes of teaching prosocial behavior as well as preventing and reducing inappropriate behavior" (pp. 7-8). Although the search profile indicates that Oliver et al. (2011) included studies published between 1950 and 2009 on classroom management and classroom organisation, the final review included only 12 studies. These were mainly published in the 1990s (9 studies), with only two published in the 1980s, and one published after 2000. The participants were both primary and secondary school students, and four studies also included special education classrooms. All selected studies were primarily focused on whole-class procedures to reduce problem behaviour such as disruptive, deviant, or aggressive classroom behaviour.

The findings revealed that teachers' classroom management practices had a significant, positive effect on decreasing problem behaviour in the classroom. Students were less disruptive and showed less inappropriate and aggressive behaviour in the treatment classrooms compared with the control classrooms. The researchers reported an effect size of Cohen's $d = 0.71$, 95% CI [0.46; 0.96] for the universal classroom management procedures. However, because only 12 studies were included, important research questions (e.g., Which components make up the most effective classroom management programs? Do differences exist between grade levels?) remained unanswered.

Another limitation of Oliver et al.'s (2011) meta-analysis is that although the procedure for the selection of studies is explained in the paper, it is unclear why most initial titles did not meet the eligibility criteria. The study started with a database of 5,134 titles, but only 94 titles were selected for further screening. Therefore, the review may not have been exhaustive. Furthermore, academic outcomes were not considered.

Similarly, the meta-analyses by Wilson, Lipsey, and Derzon (2003) and Wilson and Lipsey (2007) involved school-based interventions focused on reducing aggressive, violent, or (severely) disruptive behaviour, mainly including cognitively-oriented interventions (e.g., changing thinking or cognitive skills, social problem solving, controlling anger). Although such interventions are aimed at changing student behaviour, classroom management is usually not their explicit focus (with some exceptions such as the Good Behavior Game intervention).

2.3.3 Durlak, Weissberg, Dymnicki, Taylor, & Schellinger (2011)
Durlak et al. (2011) conducted a meta-analysis of 213 school-based, universal (school-wide) social and emotional learning (SEL) programs. These programs are aimed at enhancing students' cognitive, affective, and behavioural competencies such as self-awareness and responsible decision making. These competencies are expected to lay the foundation for better school adjustment and academic performance. SEL programs generally include some classroom management components, but this is not always the case. That is, SEL programs do not necessarily concentrate on the actions teachers take to create suitable learning environments, but concentrate on competency development among students. However, Durlak et al.'s findings are generally in line with those of Marzano et al. (2003) and Oliver et al.'s (2011) meta-analyses, reporting generally positive effects of the interventions included in the meta-analysis.

Focusing on studies that appeared in published or unpublished form before 2007, Durlak et al. (2011) selected all school-based universal studies that emphasized the development of one or more SEL skills among students from kindergarten through high school. The authors excluded studies focused on students with identified adjustment or learning problems from the analysis, as well as studies that did not focus on all students in the class. The interventions were categorized into three groups: (1) classroom-based interventions administered by regular classroom teachers, (2) classroom-based interventions administered by non-school personnel (e.g., university researchers), and (3) multicomponent programs, usually combining teacher-administered classroom interventions with a parent component, or school-wide initiatives.

Durlak et al. (2011) demonstrated that SEL programs significantly improved students' social and emotional skills, with $g = 0.57$, 95% CI [0.48; 0.67], and attitudes toward the self and others, with $g = 0.23$, 95% CI [0.16; 0.30]. Hedge's g effect sizes (at the student level) were used, which can be interpreted similarly to Cohen's d effect sizes (Cohen, 1988). Students who received SEL programs showed more positive social behaviour, $g = 0.24$, 95% CI [0.16; 0.32]. They also showed fewer conduct problems, $g = 0.22$, 95% CI [0.16 ; 0.29] and less emotional distress, $g = 0.24$, 95% CI [0.14 ; 0.35]. Moreover, the effect size for academic achievement was $g = 0.27$, 95% CI [0.15; 0.39]. A notable finding was that students' academic achievement improved significantly only when teachers implemented the intervention or when multicomponent programs were used. Implementation by non-school personnel did not yield significant results regarding students' academic achievement.

Furthermore, multicomponent program effects were comparable to those obtained using classroom-based interventions administered by teachers, although only the latter programs significantly improved SEL skills and positive social behaviour.

Analysis of a subsample of 33 studies in which follow-up tests were administered at least 6 months after the intervention ended revealed that the effects declined over time, but that they remained significant. These results indicate that SEL programs generally have lasting positive effects on students' social and emotional learning.

2.4 Frequently implemented classroom management programs

Below, five classroom management programs that are frequently implemented by primary schools[1] are described in order to illustrate the types of programs used in classrooms currently, and to make the differences between existing programs more tangible. The five programs are (1) School-Wide Positive Behavior Support, (2) Promoting Alternative Thinking Strategies, (3) The Good Behavior Game, (4) Zippy's Friends, and (5) Second Step. All except the last have a Dutch equivalent. We describe each program's aims, theoretical underpinnings, intensity, format, and effectiveness. To facilitate comparison between the different programs, as the current meta-analysis was aimed at identifying effective CMS/CMP, we classified each program in accordance with our above-described categorization system (see section 2.2 for the different types of classroom management interventions); this is provided at the end of this section to facilitate comparison between programs.

2.4.1 School-Wide Positive Behavior Support
School-wide positive behavior support (SWPBS) was developed in the USA, where over 16,000 schools, now in various stages of implementation, have adopted the program (Bradshaw, Waasdorp, & Leaf, 2012; Horner et al., 2009). It is a whole-school (and system-wide) approach, intended to create a social culture and to provide intensive behavioural support, both of which are needed for all students to achieve academic and social success. It is preventive rather than reactive, and it combines primary, secondary, and tertiary prevention measures regarding student behaviour. The primary tier involves defining, teaching, monitoring, and rewarding a small set of behavioural expectations for all students across classroom and non-classroom settings (Horner et al., 2009). Schools continually measure students' social behaviour, which permits early intervention and supports further decisions. In this way, they work on a data-driven basis. If more severe individual problems are identified or structural changes are needed, a secondary tier is brought into action. This secondary tier includes behavioural support for students 'at risk' and focuses on problem behaviour. The

---

[1] In section 3.1 we explain further why other (related) programs were not included here.

tertiary tier provides highly individualized interventions to address higher intensity problem behaviours when necessary (Horner et al., 2009).

The program is based on the principles of behaviour analysis (Anderson & Kincaid, 2005; Sugai & Horner, 2002, 2006). Schools that adopt the program are expected to set up a school-wide reward system for good behaviour, rather than punishment systems for bad behaviour (Anderson & Kincaid, 2005). Implementation of SWPBS in the USA is often initiated at state level; states also may provide personnel experienced in the training and support practices associated with the approach (Horner et al., 2009).

SWPBS is not a packaged approach, and thus schools and even departments or settings within schools may vary in the sets of rules they use, given the above-mentioned general features of the program (Anderson & Kincaid, 2005). At class level, teacher practice may, for example, typically consist of teaching expectations and target behaviours to students in classroom settings as well as in other target school environments, on the one hand, and systematically providing acknowledgment for successfully meeting those behavioural expectations, on the other hand (Solomon, Klein, Hintze, Cressey, & Peller, 2012, referring to McCurdy, Mannella, & Eldridge, 2003).

Although evaluation reports concerning SWPBS thus far are positive and show the approach to be implementable, real experimental evidence regarding its effects is just coming on stage (Chitiyo, May, & Chitiyo, 2012; Horner et al., 2009; see also Bradshaw et al., 2012; Solomon et al., 2012; Sørlie & Ogden, 2007).

In the Netherlands, the approach (SW)PBS has been introduced in primary education by 'Kenniscentrum SWPBS Nederland', Duivendrecht (www.swpbs.nl). Eight primary schools (in regular and special education) adopted SWPBS as a pilot in 2009, followed by more schools in the following years (http://wij-leren.nl/school-wide-positive-behavior-support.php). Up until now, there has been no Dutch research on the effectiveness of the program. For this reason, it is not yet recognized as effective by the 'Nederlands Jeugdinstituut' (2013), but merely regarded as a sound program; transferable and underpinned by international practice and effectiveness research.

In the present study, we classified the SWPBS program as 'teachers' behaviour focused' and 'students' behaviour focused'.

2.4.2 Promoting Alternative Thinking Strategies

The Fast Track (PATHS) intervention was developed in the USA as a universal service from an initial Fast Track selective prevention model for children at risk for behavioural problems. In this program, small-group social skills interventions are combined with academic tutoring in which parenting support classes are provided and home visits are conducted. The PATHS intervention is aimed at preventing the (further) development of violent and aggressive behaviour in children, lowering the risk of later juvenile and adult violence as well as other

social and academic maladaptive outcomes. It is mainly school-based, as schools are the only setting with almost universal access to children (Crean & Johnson, 2013).

The central component in the PATHS universal intervention is the school-based PATHS curriculum, which is a scripted curriculum in social and emotional skills taught on a regular basis throughout the school year. The PATHS curriculum contains 131 lessons in which the focus is on skills related to understanding and communicating emotions. The program aims to increase positive social behaviour, and to enable children to achieve self-control and other steps in social problem solving. The PATHS lessons may be flexibly implemented over the primary school years. In these lessons, skill concepts are presented by various means, such as direct instruction, discussion, modelling stories, and video presentations. Subsequently, the skills are practiced by pupils in discussions and role-playing activities. (For more information on the curriculum we refer to Greenberg and Kusché (2002) and Bierman, Greenberg, and the Conduct Problems Prevention Research Group (1996)). In addition to this lesson-based curriculum, the PATHS intervention emphasizes the need to implement the PATHS principles during the rest of the school day. As part of the program, school-based support for teachers as well as consultation activities with school principals are provided by the PATHS project staff (Crean & Johnson, 2013; Greenberg et al., 2010).

PATHS is based on the Affective-Behavioral-Cognitive-Dynamic (ABCD) model of development (Greenberg & Kusché, 1993). In this model, early emotional development is identified as a precursor to other ways of thinking. Moreover, the curriculum places special emphasis on neurocognitive models of development, by promoting the development of children's inhibitory control and having them verbally identify and label feelings and emotions in order to manage these (Riggs, Greenberg, Kusché, & Pentz, 2006).

In the USA, a considerable amount of research has been done on the effectiveness of the program since the nineties. In general, positive effects on students' social and emotional competence and behaviour have been found.

In the Netherlands, a Dutch version of PATHS, called PAD, was developed by 'Seminarium voor Orthopedagogiek', Utrecht, and has been introduced in regular and special primary education. The first research outcomes on the effectiveness of the program - after the program had been implemented for one year - showed moderate positive effects (Louwe, Van Overveld, Merk, De Castro, & Koops, 2007). It was recognized by the 'Nederlands Jeugdinstituut' (2013) as "proven effective by good evidence" (the second-highest category out of four).

In the present study, we classified the PATHS program as 'students' behaviour focused' and 'students' social-emotional development focused'.

2.4.3 The Good Behavior Game

The Good Behavior Game (GBG) is a classroom-based program targeting the prevention of and early intervention in aggressive and disruptive behaviour. The basic principle of the game was stated by Barrish, Saunders, and Wolf (1969); they defined the game as a "classroom behavior management technique, based on reinforcers natural to the classroom, other than teacher attention" (p. 119). The game involves competition for privileges available in almost every classroom (see Dolan, Turkkan, Werthamer-Larsson, & Kellam, 1989).

In GBG, first, appropriate behaviour is explicitly defined, and when students demonstrate such behaviour it is systematically rewarded. Appropriate behaviours are formulated as rules students have to comply with. They may be stated by way of bans, e.g., "No one is to be out of his seat without permission" (Barrish et al., 1969) or as positively formulated classroom rules, e.g., "In the classroom, we work quietly and stay in our seats" (Leflot, Van Lier, Onghena, & Colpin, 2013). The rules may differ according to the specific tasks children have to complete or the lessons being taught.

Second, GBG facilitates positive interaction between (disruptive and non-disruptive) children through a team-based approach (Van Lier, Vuijk, & Crijnen, 2005; Witvliet, Van Lier, Cuijpers, & Koot, 2009), using group contingencies. This approach divides students in each class into two or more teams, each containing students both with and without behaviour problems. The teams compete for privileges, and each team as a whole may be punished for the inadequate behaviour of its members (e.g., losing points earned by the team in the weekly contest) (Dion et al., 2011) or rewarded for helping members to comply with classroom rules (Witvliet et al., 2009). In this way, the GBG directly intervenes in the children's social context (Van Lier et al., 2005) and is supposed to bring about the positive peer interactions that underlie the effect of the program on student behaviour (Witvliet et al., 2009).

Research on GBG has demonstrated positive effects of the program on various outcome measures, varying from diminishing aggressive and disruptive behaviour, attention-deficit/hyperactivity problems, oppositional defiant problems, and conduct problems, to preventing the development of antisocial personality disorders and postponing (or preventing) tobacco use in early adolescence (Van Lier et al., 2005).

A Dutch version of GBG has been introduced in the Netherlands, the so-called 'Taakspel', developed by CED group/PI Rotterdam. Research on the effects of the program showed results comparable to those found internationally (Van der Sar, 2004). The program effectively reduces the disruptive behaviour of students. It is the only intervention program that has been recognized by the 'Nederlands Jeugdinstituut' (2014) as "proven effective by strong evidence", the highest category out of four (Spilt, Koot, & Van Lier, 2013a; Spilt, Koot, & Van Lier, 2013b; Nederlands Jeugdinstituut, 2014). For a description of the program in Dutch, see Witvliet, Van Lier, Cuijpers, and Koot (2010) and the publication of the Nederlands Jeugdinstituut (2013).

In the present study, we classified the GBG program as 'teachers' behaviour focused' and 'students' behaviour focused'.

2.4.4 Zippy's Friends

Zippy's Friends is a universal school-based preventive program for pre-school and first-grade children, to help them cope better with everyday adversities; it aims to avoid serious problems in childhood, adolescence, and adulthood. The program provides training in a variety of coping skills and in adapting those skills to different situations. It also teaches social and emotional skills that facilitate adaptive coping behaviour. Transfer to real-life situations is an important component of the program (Monkeviciené, Mishara, & Dufour, 2006).

The program usually involves 24 weekly sessions of about 50 minutes. It is built around a set of six illustrated stories about a group of children and a pet insect called Zippy. The following themes are addressed: a) understanding feelings, b) communication, c) 'making and breaking' relationships, d) conflict resolution, e) dealing with change and loss, and f) general coping skills. Children work on these topics through drawing, role playing, performing exercises, play, and dialogue (Holen, Waaktaar, Lervag, & Ystgaard, 2012). Teachers implement the program; they usually receive two days of training in advance and are supervised during the course of the program.

The program was developed as a mental health program in the late nineties, after being initiated by Befrienders International, a non-profit suicide prevention organization. By 2006, the program had been used for young school children in several countries in Europe, South America, and Asia. It is now distributed by the British 'Partnership for children' (www.partnershipforchildren.org.uk).

Available research outcomes since 2006 indicate that the program has significant effects on children's social skills and coping abilities. The program can help decrease children's problem behaviour (Holen et al., 2012; Mishara & Ystgaard, 2006). Participation in the program is related to better adaptation in the transition from kindergarten to first grade (Monkeviciené et al., 2006). Moreover, it has a positive effect on the social climate in the classroom: it reduces bullying and improves children's feeling of being good at school ('academic skills') (Holen, Waaktaar, Lervag, & Ystgaard, 2013). The program is classified as a commendable program according to the Child Welfare League of America's criteria (Dufour, Denoncourt, & Mishara, 2011).

In the Netherlands, the program is known as 'Zippy's Vrienden'; it was developed by 'Stichting Kids en Emotionele Competenties (KEC)' for primary school children aged 5 to 8 years, from Kindergarten to grade 2 (www.zippysvrienden.nl). It was recognized by the ´Nederlands Jeugdinstituut´ (2014) as well underpinned (lowest category out of four). Thus far, no research has been conducted on the effectiveness of the Dutch version of the program.

In the present study, we classified Zippy's Friends as 'students' social-emotional development focused'.

2.4.5 Second Step

Second Step is a universal school-based social-emotional intervention program for children ranging from preschool/Kindergarten to the 8th grade. Central to the program is the understanding that behaviour depends on social cognition (goals, beliefs) and emotions, in addition to behavioural and cognitive skills. The program's general goal is to prevent aggressive behaviour by increasing prosocial behaviour (Grossman et al., 1997). Specifically for the school context, the program is aimed at preventing maladaptive classroom behaviour that might compromise the learning environment and cause conduct problems later on in children's school careers and adult lives.

Second Step contains a developmentally sequenced set of activities in the classroom, supported by commercially available curriculum materials, staff training, and staff training materials. Classroom teachers implement 25-40-minute scripted lessons, in which key concepts are introduced by presenting visual materials. Students are then stimulated to develop thoughts, feelings, and behaviours, to practice perspective-taking, and to develop specific strategies in response to the illustrated situations. Children practice a variety of behaviours through role playing in the classroom. The curriculum of 51 lessons covers three thematic units: empathy training, impulse control and problem solving, and anger management (Frey, Nolen, Edstrom, & Hirschstein, 2005; Schick & Cierpka, 2005).

Second Step was developed in the eighties by the American 'Committee for children' as a violence prevention program (http://www.cfchildren.org/second-step.aspx). Evaluation studies showed that the program reduced children's aggression and problem behaviour, and that it fostered prosocial and neutral interactions and social competence (Frey et al., 2005; Grossman et al., 1997; Holsen, Smith, & Frey, 2008; McMahon, Washburn, Felix, Yakin, & Childrey, 2000; Schick & Cierpka, 2005). Some results, however, are contradictory, and effects may vary depending on classroom and gender. The program has been translated into German and Norwegian (see Holsen et al., 2008; Schick & Cierpka, 2005). To our knowledge, no Dutch version of the program is available.

Finally, we would like to explain why, although we excluded training programs that were primarily focused on social skills (see paragraph 3.1) from our study, we decided to present the Second Step program here. Second Step does indeed focus on the development of social skills, but does so as part of a broader intervention program focusing on enhancing student behaviour and students' general social-emotional development. Training in social skills is increasingly seen as an important part of the primary school curriculum; however, it can be seen as a relevant component of students' social-emotional development rather than a separate set of skills.

In the present study, we classified the Second Step program as 'students' behaviour focused' and 'students' social-emotional development focused'.

In sum, the five programs discussed above differ in some respects regarding their main focus.[2] The SWPBS and GBG programs can be considered both 'teachers' behaviour focused' and 'students' behaviour focused', whereas the PATHS and Second Step programs can be considered 'students' behaviour focused' and 'students' social-emotional development focused'. Finally, Zippy's Friends can be considered 'students' social-emotional development focused'. We therefore conclude that all programs have at least one student-focused component in their intervention, but only two contain teacher-focused components (e.g., improving teachers' use of classroom rules and procedures). Remarkably, although the importance of establishing positive teacher-student relationships (our second classification) is emphasized in all programs, in none of the programs is this component explicitly integrated in the intervention, or at least not in the descriptions of the interventions.

---

[2] We refer to these programs as classroom management programs, however, we acknowledge that not all programs have presented themselves in such terms. The inclusion of these programs follows the broad definition of classroom management of Evertson and Weinstein (2006).

3   Method

3.1 Literature search

In the literature search, which was aimed at identifying studies in which the effectiveness of classroom management programs and their accompanying strategies was investigated, we included the online databases ERIC, Web of Science, PsycINFO, and Picarta from 2003 until 2013; we focused on peer-reviewed journal articles and abstract collections. The keyword searches included the following terms: *classroom management, classroom organisation/organization, behaviour(al) management, classroom technique(s), teacher/teaching strategy/strategies, classroom discipline, group contingency/contingencies.* These keywords were combined with: *academic outcomes, academic achievement, performance, on-task/off-task/time-on-task, student engagement, academic engagement, student behaviour, classroom behaviour.* Both British English and American English spelling were used. Studies that considered grades 1 to 6, elementary education, primary education, preschool education, kindergarten, and early childhood education were included. Additionally, the journals 'Teaching and Teacher Education' and 'Pedagogische Studiën' were consulted for relevant references, as were the publications of Hattie (2009) and Evertson and Weinstein (2006).

After the first round, specific classroom management intervention programs were used as additional search terms. The selection of those interventions was based on the results of the first round (the programs that were identified in this round were *The Good Behavior Game, The Color Wheel System,* and *Classroom Organization and Management Program).* Moreover, we found the study by Freiberg and Lapointe (2006), who listed numerous behavioural intervention programs in American education. From this overview, we selected the programs that focused on the entire classroom and used students' behaviour or achievement as outcome measures (the programs identified in this step were the *Daily Behavior Report Card, Peacebuilders, Promoting Alternative Thinking Strategies,* and *School-Wide Positive Behavior Support*). Through the Best Evidence Encyclopaedia, we found one additional program focused on classroom management, *Consistency Management & Cooperative Discipline*, which was also included in the additional literature search.

The reference lists of the selected papers were then checked for publications that we had not found in the previous steps. Some of these publications referred to another relevant classroom management intervention program, *Zippy's Friends;* we decided to use this search term in the databases to find related papers. Finally, we decided to include two new search term combinations, *social-emotional learning* and *social-emotional outcomes* in combination with *school*, because we discovered that some of the interventions we had selected used these terms to explain the content of their intervention (e.g., PATHS). In social and emotional

learning programs, students "develop skills to recognize and manage their emotions, develop caring and concern for others, make responsible decisions, establish positive relationships, and handle challenging situations effectively" (Weissberg, Resnik, Payton, & O'Brien, 2003, pp. 46-47).

Although several school-wide programs focused on anti-bullying include teacher strategies to reduce problem behaviour in class, studies aimed at investigating these programs were excluded from the present study. Several reviews specifically focusing on this topic have already been conducted in recent years (e.g., Ferguson, Miguel, Kilburn, & Sanchez, 2007; Merrell, Gueldner, Ross, & Isava, 2008; Ttofi & Farrington, 2011). For anti-bullying programs that have been successfully implemented in Dutch primary education, we would like to point out the KiVa program (for details on the program see Kärnä, Voeten, Little, Poskiparta, Kaljonen, & Salmivalli, 2011) and the PRIMA method (see Van Dorst, Wiefferink, Dusseldorp, Galindo Garre, Crone, & Paulussen, 2008). Similarly, training programs primarily focused on social skills were excluded, because these generally concentrate on enhancing students' mental resilience rather than their general social-emotional development (e.g., developing empathy). However, when training in social skills was part of another program that met our inclusion criteria, the studies were included in the meta-analysis.

3.2 Inclusion criteria

The studies had to meet the following criteria to be eligible for inclusion:
1) The focus of the study was on CMS of teachers or CMP implemented by teachers in regular, primary school classrooms.
2) The interventions needed to focus on (basically) all students in the classroom, i.e., interventions aimed at changing individuals' or small groups' behaviour were not eligible.
3) The outcome variable had to include measures of academic outcomes, behavioural outcomes, social-emotional outcomes, motivational outcomes, or other relevant student outcomes (e.g., time-on-task, self-efficacy, peer acceptance).
4) The studies had to be (quasi-)experimental designs with control groups (no treatment or treatment as usual). They had to meet at least one of the following criteria: (a) participants were randomly assigned to treatment and control or comparison conditions, (b) participants were matched into treatment and control conditions and the matching variables included a pretest for the outcome variable, or pretest differences were statistically controlled for using ANCOVA, (c) if subjects were not randomly assigned or matched, the study needed to have a pre-posttest design with sufficient statistical information to derive an effect size or to estimate group equivalence from statistical significance tests.

After the initial screening of the titles and abstracts to eliminate off-topic papers, 241 studies were selected for further inspection. These studies were divided among three researchers to determine whether they met the inclusion criteria. A second selection round using the four stated criteria was conducted to determine which studies met the inclusion criteria. In this selection round, all studies were initially categorized into three groups: eligible, possibly eligible, not eligible. The researchers met on several occasions to discuss how stringent the inclusion criteria needed to be (e.g., whether Kindergarten classrooms are part of primary schools or not. We decided that the studies conducted in these classrooms were eligible for inclusion). Further, all studies that were initially labelled 'possibly eligible' were discussed by the three researchers involved in this selection process. When necessary, a second researcher read the study. Moreover, all studies that were labelled 'not eligible' or 'eligible' were checked by a second researcher, based on the abstract, and the full paper when the abstract did not provide enough details. The final decisions for inclusion ('eligible') were thus based on complete consensus. Following this procedure, 47 studies were selected for the meta-analysis.

The main reasons for excluding 194 studies followed from the inclusion criteria. Most (135 studies; 70%) did not have a suitable research design (no control group, correlational studies, no empirical data) and therefore did not meet criterion 4. Moreover, 21 studies did not focus on classroom management at all, or were not conducted in regular, primary school classrooms (criterion 1). In 10 studies, the intervention was not focused on all students in the classroom (criterion 2), and 7 studies did not include relevant student outcome variables, but, for example, only outcome variables at the school level (e.g., retention rates) (criterion 3). For 21 studies, there were other reasons for exclusion: mainly because not enough statistical data were provided to compute effect sizes or the datasets of studies we had already included were used without new relevant additional outcome measures.

3.3 Coding of the studies

The 47 selected studies were coded for further investigation, initially including the following information: CMS/CMP under investigation (teachers' behaviour focused, teacher-student relationship focused, students' behaviour focused, students' social-emotional development focused), duration of the intervention, number of intervention sessions, outcome variables (student performance [reading, writing, arithmetic, science, other], time-on-task, student behaviour, student engagement), sample characteristics (average students, learning problems, behavioural problems, low SES, high SES, grade level, age), country where the study was conducted, educational context (during instruction, independent seatwork, cooperative learning, lesson transitions), classroom setting (group settings, frontal placement, thus facing the teacher), research measurement instrument (designed by authors, unstandardized instrument designed by others, standardized instrument designed by others), design (pre-posttest, control group, random sample), sample size, the reported effects, the number of

schools or classes included, and whether the data were reported at the student, class, or school level.

To code the CMS/CMP under investigation, inter-rater reliabilities were calculated. Two researchers showed 89% agreement (46 studies,[3] 4 categories), resulting in an inter-rater reliability (Cohen's Kappa) of 0.78. The differences in coding concerned 13 studies. In 9 cases, one of the researchers had indicated more categories than the other; we decided to combine the scores of the two researchers. For the four remaining studies, the coding differences were more substantial. Both researchers reread these articles and changed their initial coding where they thought necessary. This resulted in two studies on which the researchers agreed, and two studies in which their scores were combined (as described above).

We were also interested in the effectiveness of frequently used CMP. Therefore, after the initial coding, the studies were categorized into groups with the same intervention (a minimum of 3 studies per intervention): (1) other, (2) School-Wide Positive Behavior Support [SW-PBS], (3) Promoting Alternative Thinking Strategies [PATHS], (4) Good Behavior Game [GBG], (5) Second Step, and (6) Zippy's Friends. The duration of the intervention was categorized into three groups: less than 13 weeks, between 13 weeks and 1 year, and longer than one year. Dichotomous variables were added to indicate whether the study was conducted in the USA or in a different country (studies conducted in the USA were largely overrepresented). We included a variable indicating whether participants – students, classes/teachers, or schools – were randomly assigned to intervention and control groups. The outcome measures were recoded into academic outcomes, behaviour (including self-control), social-emotional outcomes (including social skills, social competence, emotion recognition, coping, and empathy), motivation, and other outcomes (e.g., time-on-task, self-efficacy, peer acceptance). Outcomes measured using highly unreliable instruments were not included. An additional categorical variable indicated whether the outcome measures were rated by the students (self-rating), by their teachers, by a researcher/observer, or by other people, usually parents or peers. We decided to include only those that were rated by the students themselves, the teachers, or the researchers/observers, because we considered it to be more difficult for parents and peers to assess the students' behaviour in the classroom only, without taking behaviour outside the classroom into account. The socioeconomic status of the students was recoded into (0) more than 40% free or reduced lunch (low SES) or (1) less than 40% free or reduced lunch (medium or high SES). The grade levels included in the studies were categorized into (0) both lower and higher grades, (1) pre-K to grade 1, and (2) grade 2 and up. Finally, a dichotomous variable indicated whether regular students or the students with frequent problem behaviour were assessed (that is, despite the fact that the intervention was focused on the entire class).

---

[3] The studies of Holen, Waaktaar, Lervåg, and Ystgaard (2012) and Holen, Waaktaar, Lervåg, and Ystgaard (2013) are counted as one study, because they report the results of the same intervention study. The only difference between the studies is the outcome variables reported.

Originally, we were interested in the results of the programs in follow-up tests. Follow-up tests were used in only two out of the selected 47 studies, hence this variable could not be taken into account during the analyses. We were also interested in the educational context (e.g., whether it concerned instruction, independent seatwork, cooperative learning, or lesson transitions), but in most cases the intervention was implemented throughout the day rather than in specific educational contexts, or the educational context was not reported. Hence, this study characteristic could not be taken into account either. Furthermore, we aimed to include the classroom setting (group settings or frontal placement), but only four studies reported this information.

3.4 Data analysis

Meta-analyses were performed using the program Comprehensive Meta-Analysis of Biostat (CMA; Borenstein, Hedges, Higgings, & Rothstein, 2009). In a meta-analysis, the unit of analysis is not the individual participant, but the effect size determined on the basis of primary studies' outcomes. Therefore, an important part of the analyses is (re)calculation of the effect sizes, to enable a useful comparison between the reported effects of the different studies. In most of the intervention studies, the results data were based on a pretest-posttest control group design. Using the above-mentioned program, Hedges' g was calculated, which is the adjusted standardized mean difference (d) between two groups, based on the pooled standard deviations. Hedges' g is particularly useful for a meta-analysis of studies with different sample sizes. We defined the direction of the effect in such a way that a positive effect size indicates that the intervention group did better than the control group (e.g., higher academic performance, better behaviour), and a negative effect indicates that the control group did better than the intervention group. We defined the effects at the level of the students, and not at the level of the class or school. Most of the data in the primary studies were also reported at the level of the students, but in 4.6% of the reported data the class or school was the unit of analysis. In these cases, we recomputed the class/school-level effect sizes by multiplying them by the square root of the intraclass correlation, as Hedges (2007) prescribes. Hedges and Hedberg (2007) reported an average intraclass correlation value (in models that corrected for pretest scores) of about 0.1 for primary school students' performance in reading and mathematics. Average intraclass correlations for non-academic outcomes were not reported. Our meta-study, however, included two studies (Benner, Nelson, Sanders, & Ralston, 2012; Raver, Jones, Li-Grining, Zhai, Metzger, & Solomon, 2009) in which these were reported for behavioural outcomes, and both came to an average value of 0.1. This is also the value used in the meta-analyses of What Works Clearinghouse (2014). Therefore, we used 0.1 as the intraclass correlation value for all our recomputations. The guidelines suggested by Cohen (1988) state that an effect size of $d = 0.2$ can be interpreted as a *small* effect, $d = 0.5$ as a

medium effect, and *d* = 0.8 as a *large* effect. This interpretation also applies to Hedges' *g* effect size measures.

CMA was also used to compute the variances of the individual interventions' effect sizes. This information was used to perform weighted analyses. The weight assigned to each intervention is the inverse of the variance. In this way, interventions with lower variances (which were the interventions with larger sample sizes) had a greater effect on the calculated summary effects.

The summary effects were estimated using a random effects model. Moderator analyses (with ANOVA for meta-analytical data) were conducted using a mixed-effects model. In the analysis, the coded characteristics of CMS/CMP were modelled as predictors of the differences between the effects found. The predictors were categorized at the level of the intervention, and the dependent variables were the sizes of the effects (for all student outcomes) of these interventions.

An elegant feature of CMA is that it is possible to examine the probability of biased results due to a phenomenon called *publication bias*. Studies are more likely to be published when the effects found in the study are significant, or when the study is based on a large sample size. Studies based on smaller sample sizes and reporting no significant effects might, therefore, be underrepresented in the meta-analysis. CMA is used to analyse the relationship between sample size and effect size. The program assumes that, if there is a relationship between the two constructs, this can be attributed to missing studies. Furthermore, it estimates to what extent the results of the meta-analysis are likely to be biased.

4   Results

4.1 Characteristics of the intervention studies

We first present the descriptive characteristics of the selected studies. The results of 46 studies[4] were used in the analyses, which together report the findings of 54 intervention studies. Table 1 gives an overview of the characteristics of the intervention studies.

The focus of most of the intervention studies was on changing the students' (students' behaviour and/or students' social-emotional development) and/or the teachers' (i.e., their CMS) behaviour through long-term interventions; the shortest intervention lasted 6 weeks and the longest three years. Only two intervention studies were explicitly focused on changing teacher-student relationships. A large variety of interventions was implemented in the studies. Only the PATHS program was implemented relatively often, in 10 intervention studies.

About three-quarters of the intervention studies were conducted in the USA; the other studies were mainly conducted in European countries (Norway, Luxembourg, Belgium, the Netherlands, Germany, Denmark, Lithuania, Turkey, UK) and in Canada. Regarding the student sample characteristics, we found that both lower and higher grade levels were represented in the selected intervention studies and that regular students (without serious behaviour problems) were commonly included. Although the socioeconomic status of the students was not indicated in several studies, we found that low-SES students were overrepresented in the selected studies compared with mid- and high-SES students. Three intervention studies reported results for boys and girls separately, and one intervention targeted boys only. The other 50 interventions did not distinguish their results according to students' sex.

Results were often reported for more than one outcome type. Table 1 shows how often each outcome was reported in total in our sample of interventions. Student behaviour was by far the most common student outcome (44%), followed by social-emotional outcomes (28%) and academic outcomes (17%). In a few studies, student motivation (6%) or another outcome measure at the student level (5%; e.g., social integration, peer acceptance, self-confidence) was reported. Also, intervention effects were often estimated using more than one measurement instrument. The total number of tests used in the interventions was 262. In half of these, the teachers rated the student outcomes, and in one third of the tests student self-reports were used. In a few cases, an external observer rated the student outcomes.

---

[4] The studies by Holen, Waaktaar, Lervåg, and Ystgaard (2012) and Holen, Waaktaar, Lervåg, and Ystgaard (2013) are counted as one study, because they report the results of the same intervention study. The only difference between the studies is the outcome variables reported.

Table 1 *Overview of the characteristics of the 54 intervention studies*

		# interventions	% interventions
Duration of the intervention:	< 13 weeks	6	11.1
	13 weeks to 1 year	30	55.6
	> 1 year	18	33.3
Focus of the intervention:	teachers' behaviour	29	53.7
	students' behaviour	46	85.2
	students' social-emotional development	40	74.1
	teacher-student relationship	2	3.7
Name of the intervention:	School-Wide Positive Behavior Support	3	5.6
	Promoting Alternative Thinking Strategies	10	18.5
	Good Behavior Game	4	7.4
	Second Step	3	5.6
	Zippy's Friends	3	5.6
	other	31	57.4
Country:	USA	39	72.2
	other	15	27.8
Grade years:	pre-K and grade 1	22	40.7
	grade 2 − 6	20	37.0
	both	12	22.2
Type of student sample:	regular students	46	85.2
	students with behaviour problems	5	9.3
	missing	3	5.6
Sex:	girls	3	5.6
	boys	4	7.4
	no results specification for students' sex	50	92.6
Socioeconomic status:	low SES	27	50.0
	mid and high SES	15	27.8
	missing	12	22.2
Outcome variables: (an intervention can have more than 1 outcome type):	academic outcomes	17 (37 tests)	17.3
	behaviour outcomes	43 (147 tests)	43.9
	social-emotional outcomes	27 (58 tests)	27.6
	motivational outcomes	6 (10 tests)	6.1
	other outcomes	5 (10 tests)	5.1
Rater (total tests = 262):	teacher	137	52.3
	student	89	34.0
	observer	36	13.7

4.2 Effects of the interventions

The findings of meta-analytical analysis show that the classroom management interventions have a small but significant effect on various student outcome measures. Table 2 reports the statistics for all outcomes together, and for each outcome separately. These statistics are indices of the average effect sizes (Hedges' *g*), their variation (*SE*), and the source of variation: true differences or random error (I^2). The Q-statistics for the outcomes show if there is significant heterogeneity among the effect sizes. If so, it is likely that the interventions do not share the same true effect size. For the overall outcome, the Q-statistic indicates that this is the case, suggesting that the variations in effect size reflect real differences between the interventions. I^2 indicates the percentage of the heterogeneity in intervention effect sizes that can be explained by differences between the interventions. In Table 2 below, one can see that I^2 for the overall effect was 84.52, which suggests that 84.52% of the dispersion of the interventions' effect sizes reflects real differences in effect size, and that 15.48% is due to random error. This also applies to each of the outcomes separately. T^2 is the estimated population variance of the effect sizes.

In an additional analysis, we examined whether the effect sizes differed significantly between the various groups of outcomes, but they did not (*Q*-between = 5.29; *df* = 4; *p* = 0.26).

Table 2 *Effects of classroom management interventions*

Outcome	Hedges' *g* (*SE*)	Q (*df; p*)	I^2	T^2
Overall	0.22 (0.02)**	342.45 (53; 0.00)**	84.52	0.01
Academic	0.17 (0.04)**	64.71 (16; 0.00)**	75.28	0.01
Behaviour	0.24 (0.03)**	183.55 (42; 0.00)**	77.12	0.02
Social-emotional	0.21 (0.03)**	117.23 (26; 0.00)**	77.82	0.02
Motivation	0.08 (0.08)	16.00 (5; 0.01)*	68.74	0.02
Other	0.26 (0.10)*	11.08 (4; 0.03)*	63.90	0.03

**p < 0.01; * *p* < 0.05

The findings furthermore revealed that the meta-analysis was subject to some publication bias. Duval and Tweedie's *Trim and Fill* method (Borenstein et al., 2009; Peters, Sutton, Jones, Abrams, & Rushton, 2008) for a random effects model showed that, for all outcomes together, the meta-analysis lacked 12 interventions on the left side of the mean; this is a lower effect size than average. If these 12 interventions had been added, the average effect size would have been slightly lower with Hedges' *g* = 0.17 (*SE* = 0.02). We also found publication bias for each outcome separately, except for the motivational outcomes. Duval and Tweedie's method indicated that, for the academic, social-emotional, and 'other outcomes' (e.g., time-on-task, self-efficacy, peer acceptance), interventions with lower effect sizes were lacking, and for the behavioural outcomes one intervention with a higher effect size was lacking. Figure 1 shows

the funnel plots of the relationship between standard error and effect size for all outcomes together and for each outcome separately. The figures display the observed and imputed interventions. The imputed interventions are those that were estimated as probably lacking due to publication bias. The interventions with a small sample size generally have a larger standard error and appear at the bottom of the figure.

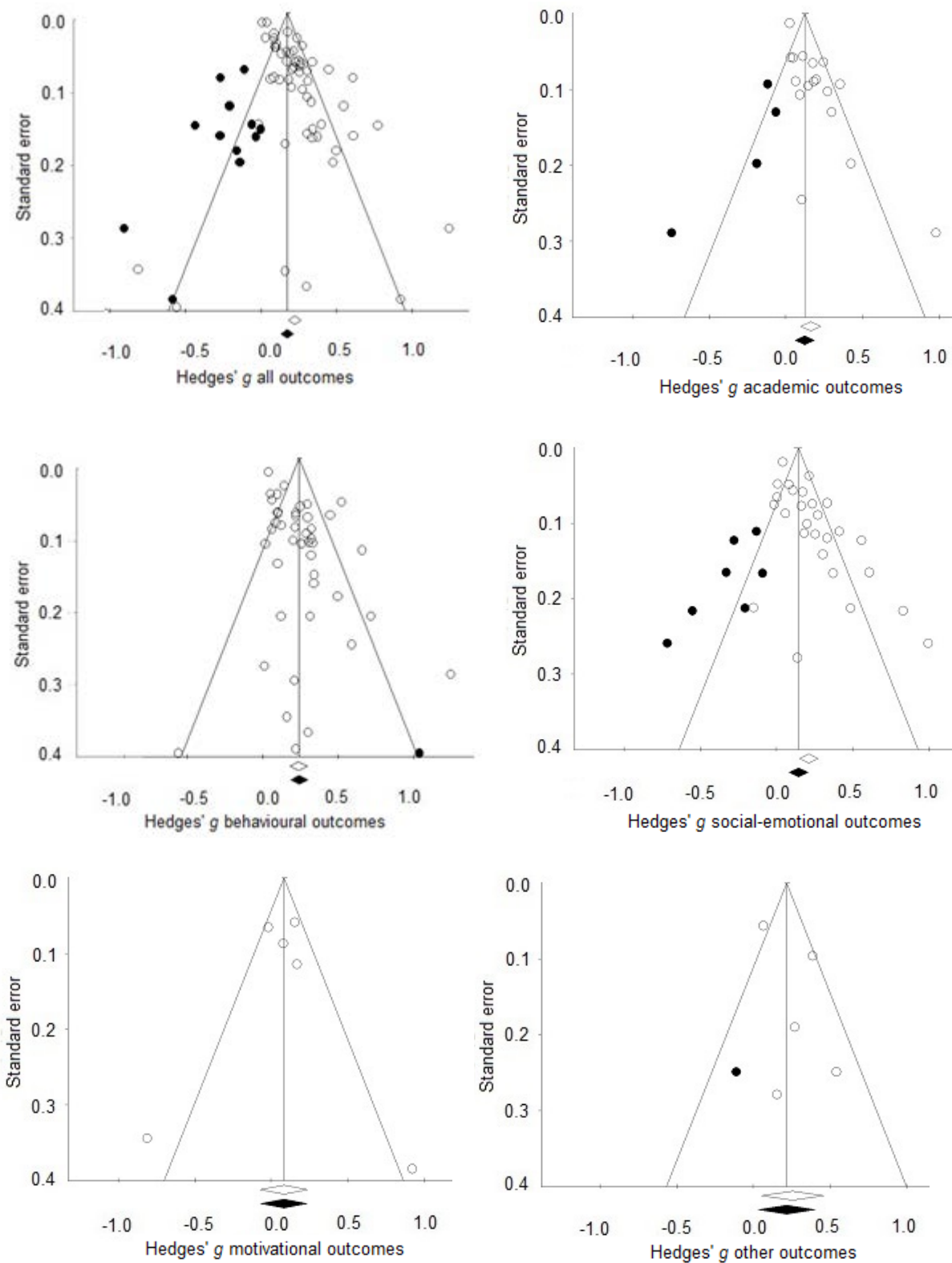Figure 1. *Funnel plots of standard error by effect size for the interventions. The observed interventions are represented by an open circle; imputed interventions are represented by a filled circle. The diamonds at the bottom represent the summary effect and its confidence interval; the open diamond for the observed interventions only, and the filled diamond for the observed and imputed interventions.*

4.3 Moderator analyses

We examined the relationship between the intervention effects and the type of classroom management intervention. Table 3 reports the average effects for each component of the interventions that we distinguished based on their focus. The table presents the estimated effects for interventions that include a particular component ('component included') and for interventions that do not include a particular component ('component not included'). As interventions can focus on multiple components at once, we also examined whether the effectiveness of the intervention depended on the number of components it addressed. Table 4 reports these results. In addition, Table 5 shows the effects for all the combinations of components that were present in our meta-analysis, to indicate whether a particular combination of certain components is more effective. Last, Table 6 shows the effects for five specific intervention programs of which our meta-analysis included at least three studies, and a sixth category containing the other interventions. Using meta-ANOVA, we tested for each outcome separately whether the differences in effects were significant. The Q-betweens (which follow the same logic as an F-value in regular ANOVA) are reported in the last columns of the tables.

Table 3 shows that, for all outcome types together, interventions were not more effective when they focused on changing the teachers' behaviour (e.g., keeping order, introducing rules and procedures), changing student behaviour (either students' behaviour or students' social-emotional development, or both), or improving the teacher-student relationship. However, with a p-value of exactly 0.05, the results do suggest that focusing on the social-emotional development of students had an effect. Programs that addressed this component had a slightly higher effect size than programs that did not. Taking a closer look at the different types of outcomes, it can be seen that particularly the social-emotional outcomes (e.g., empathy for other children's feelings) benefitted from programs designed to enhance students' social-emotional development. Furthermore, we found that academic outcomes seemed to benefit from a program focused on improving teachers' classroom management and their behaviour; here, the p-value was again exactly 0.05. The category 'other outcomes' showed positive effects for teacher-focused and students' behaviour-focused programs, but these results were based on very few interventions and should, therefore, be interpreted with care.

Table 4 indicates that academic outcomes were higher when interventions were focused on three or all components. The category 'other outcomes' showed higher effects for interventions with at least two components. The number of components had no effect on the remaining outcome types. As shown in Table 5, we analysed the differences between the various combinations of focus components in two ways: based on all categories and based on the categories with three or more interventions (the restricted meta-ANOVA). The latter analysis has the advantage that the number of groups in the analysis is more in line with the number of interventions included. According to Borenstein et al. (2009), a meta-analysis

should include no more than one group per approximately 10 interventions. As such, the results of the restricted meta-ANOVA results are to be preferred to the results of the analysis based on all categories. The results of the restricted meta-ANOVA suggest that none of the combinations of focus components of interventions make a difference. An interesting descriptive finding was that the most common combinations of classroom management components were programs combining a focus on students' behaviour and students' social-emotional development (18 studies), and programs combining these two student components with a teacher focus (13 studies). Slightly less common were programs which combined a focus on students' behaviour and a focus on teachers (11 studies). Other combinations of components were less frequently observed (5 different combinations across 12 studies).

Table 6 reveals that there were differences in effectiveness between the specific programs, except for the behavioural and motivational outcomes. When we focused on all outcome types together, we found all programs to have small to moderate effects, but only SWPBS had no effect. The specific programs seemed less effective than the category 'other interventions' for academic and 'other' outcomes. PATHS was found to have the highest effect on social-emotional outcomes, and SWPBS the lowest. Again, though, the results should be interpreted with care, as some averages are based on very few intervention studies.

Table 3 *Average effects (Hedges' g (SE)) for each focus component on the various outcome types*

Focus component	Component included	Component not included	Meta-ANOVA *Q-between (df; p)*
	All outcomes		
Teacher's behaviour	0.20 (0.03)**	0.24 (0.03)**	0.88 (1; 0.35)
Students' behaviour	0.21 (0.03)**	0.26 (0.05)	0.85 (1; 0.36)
Students' soc.-em. development	0.24 (0.02)**	0.15 (0.04)**	3.83 (1; 0.05) [c]
Teacher-student relationship	0.13 (0.09) [b]	0.22 (0.02)**	1.05 (1; 0.31)
	Academic outcomes		
Teacher's behaviour	0.21 (0.05)**	0.09 (0.03)** [a]	3.84 (1; 0.05) [c]
Students' behaviour	0.18 (0.04)**	0.11 (0.06) [a]	0.86 (1; 0.35)
Students' soc.-em. development	0.17 (0.03)**	0.15 (0.08)*	0.06 (1; 0.82)
Teacher-student relationship	0.24 (0.09)** [b]	0.16 (0.04)**	0.84 (1; 0.36)
	Behavioural outcomes		
Teacher's behaviour	0.21 (0.04)**	0.28 (0.04)**	1.46 (1; 0.23)
Students' behaviour	0.23 (0.03)**	0.28 (0.10)**	0.24 (1; 0.63)
Students' soc.-em. development	0.25 (0.03)**	0.20 (0.06)**	0.71 (1; 0.40)
Teacher-student relationship	0.06 (0.10) [b]	0.24 (0.03)**	2.92 (1; 0.09)
	Social-emotional outcomes		
Teacher's behaviour	0.16 (0.05)**	0.24 (0.04)	1.92 (1; 0.17)
Students' behaviour	0.20 (0.04)**	0.25 (0.05)**	0.64 (1; 0.42)
Students' soc.-em. development	0.25 (0.03)**	0.04 (0.02)* [a]	30.35 (1; 0.00)**
Teacher-student relationship	0.06 (0.09) [b]	0.22 (0.03)**	2.99 (1; 0.08)
	Motivational outcomes		
Teacher's behaviour	0.08 (0.09) [b]	0.08 (0.11)	0.00 (1; 0.98)
Students' behaviour	0.08 (0.08)	-	-
Students' soc.-em. development	0.14 (0.05)** [a]	0.01 (0.37) [a]	0.12 (1; 0.73)
Teacher-student relationship	0.08 (0.09) [b]	0.08 (0.11)	0.00 (1; 0.98)
	Other outcomes		
Teacher's behaviour	0.38 (0.08)** [a]	0.07 (0.06) [b]	10.23 (1; 0.00)**
Students' behaviour	0.39 (0.09)** [a]	0.09 (0.06) [b]	7.67 (1; 0.01)**
Students' soc.-em. development	0.18 (0.10) [a]	0.39 (0.10)** [b]	2.39 (1; 0.12)
Teacher-student relationship	0.27 (0.19) [b]	0.26 (0.12)* [a]	0.00 (1; 0.95)

**$p < 0.01$; * $p < 0.05$; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions; [c] $p = 0.05$.

Table 4 *Average effects (Hedges' g (SE)) for number of focus components*

Outcome	1 component	2 components	3 or 4 components	Meta-ANOVA *Q-between (df; p)*
Overall	0.17 (0.07)**	0.24 (0.03)**	0.20 (0.04)**	1.19 (2; 0.55)
Academic	0.11 (0.04)* [a]	0.10 (0.06)	0.23 (0.03)**	7.35 (2; 0.03)*
Behaviour	0.27 (0.10)**	0.24 (0.04)**	0.20 (0.04)**	0.76 (2; 0.68)
Social-emotional	0.16 (0.06)**	0.27 (0.05)**	0.16 (0.08)*	3.06 (2; 0.22)
Motivation	0.01 (0.37) [a]	0.16 (0.05)** [b]	0.08 (0.09) [b]	0.72 (2; 0.70)
Other	0.07 (0.06) [b]	0.36 (0.09)** [b]	0.37 (0.15)* [b]	9.70 (2; 0.01)**

**$p < 0.01$; * $p < 0.05$; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions.

Table 5 *Average effects (SE) for focus of classroom management interventions*

Focus	All outcomes	Academic outcomes	Behavioural outcomes	Soc.-em. Outcomes	Motivational outcomes	Other outcomes
Stud. Behaviour	0.06 (0.37) [a]	0.18 (0.07)** [b]	0.21 (0.07) [**b]	0.01 (0.07) [b]	0.01 (0.37) [a]	-
Stud. soc.-em. (incl. Zippy's friends)	0.20 (0.05)** [a]	0.08 (0.04) [b]	0.28 (0.13)* [a]	0.19 (0.06)** [a]	-	0.07 (0.06) [b]
Stud. behaviour + stud soc.-em. (incl. PATHS and Second step)	0.27 (0.04)**	0.05 (0.06) [b]	0.27 (0.05)**	0.29 (0.04)**	0.16 (0.05)** [b]	0.16 (0.28) [b]
Teacher + stud. behaviour (incl. SWPBS and GBG)	0.16 (0.04)**	0.16 (0.10) [a]	0.19 (0.07)**	0.04 (0.02)* [b]	-	0.39 (0.10)** [b]
Teacher + stud. soc.-em.	0.37 (0.09)** [a]	-	0.29 (0.10)** [b]	0.36 (0.10)** [a]	-	-
Teacher + stud. behaviour + stud. soc.-em.	0.20 (0.04)**	0.23 (0.04)**	0.22 (0.05)**	0.18 (0.10)	-	0.54 (0.25)* [b]
Teacher + relation + soc.-em.	0.31 (0.20) [b]	0.43 (0.20)* [b]	0.21 (0.21) [b]	-	-	0.27 (0.19) [b]
All components	0.09 (0.10) [b]	0.20 (0.09)* [b]	0.02 (0.12) [b]	0.06 (0.09) [b]	0.08 (0.09) [b]	-
Meta-ANOVA Q-between (df; p)	8.76 (7; 0.27)	13.34 (6; 0.04)*	5.09 (7; 0.65)	42.98 (6; 0.00)**	0.72 (2; 0.70)	11.08 (4; 0.03)*
Meta-ANOVA Q-between (df; p) restricted	6.85 (5; 0.23)	0.44 (1; 0.51)	1.27 (3; 0.74)	3.55 (3; 0.32)	-	-

**$p < 0.01$; * $p < 0.05$; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions. Restricted Meta-ANOVA is based on the cells with 3 or more interventions.

Table 6 *Average effects (SE) for specific classroom management interventions*

Outcome	SWPBS	PATHS	GBG	Second step	Zippy's friends	Other	Meta-ANOVA Q-between (df; p)
Overall	0.03 (0.02) [a]	0.29 (0.05)**	0.22 (0.09)* [a]	0.21 (0.05)** [a]	0.19 (0.08)* [a]	0.23 (0.03)**	46.32 (5; 0.00)**
Academic	0.01 (0.01) [b]	-	0.09 (0.11) [b]	-	0.11 (0.06)* [b]	0.19 (0.04)**	24.08 (3; 0.00)**
Behaviour	0.16 (0.13) [b]	0.26 (0.06)**	0.25 (0.09)** [a]	0.22 (0.05)** [a]	0.18 (0.08)* [a]	0.25 (0.05)**	1.24 (5; 0.94)
Social-emotional	0.04 (0.02)* [b]	0.32 (0.05)**	-	0.16 (0.07)* [b]	0.22 (0.09)* [a]	0.20 (0.05)**	34.28 (4; 0.00)**
Motivation	-	0.17 (0.11) [b]	-	-	-	0.07 (0.10)	0.52 (1; 0.47)
Other	-	-	-	0.16 (0.28) [b]	0.07 (0.06) [b]	0.38 (0.08)** [a]	10.32 (2; 0.01)**

**$p < 0.01$; * $p < 0.05$; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions.

The next moderator analyses were focused on differences related to student characteristics. Table 7 reports the statistics for sex, Table 8 for grade year, Table 9 for socioeconomic status, Table 10 for student behaviour, and Table 11 for country. None of the reported student characteristics were found to cause differences in the intervention effects. We only found a difference for socioeconomic status and for country on 'other outcomes', but these analyses were based on a very small number of interventions and should be interpreted with care.

Table 7 *Average effects (SE) for sex*

Outcome	Girls	Boys	Meta-ANOVA Q-between (*df; p*)
Overall	0.10 (0.09) [a]	0.23 (0.11)* [a]	0.88 (1; 0.35)
Academic	0.06 (0.09) [b]	0.54 (0.32) [b]	2.17 (1; 0.14)
Behaviour	0.30 (0.09)** [a]	0.19 (0.09)* [a]	0.87 (1; 0.35)
Social-emotional	-0.01 (0.10) [b]	0.02 (0.09) [b]	0.03 (1; 0.86)
Motivation	0.02 (0.10) [b]	-0.05 (0.09) [b]	0.24 (1; 0.63)
Other	0.23 (0.28) [b]	0.31 (0.26) [b]	0.05 (1; 0.82)

**p < 0.01; * *p* < 0.05; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions.

Table 8 *Average effects (SE) for grade year*

Outcome	Pre-K & grade 1	Grades 2 − 6	Both	Meta-ANOVA Q-between (*df; p*)
Overall	0.28 (0.04)**	0.17 (0.03)**	0.20 (0.04)**	4.33 (2; 0.12)
Academic	0.23 (0.04)**	0.15 (0.06)**	0.09 (0.06) [a]	4.45 (2; 0.11)
Behaviour	0.27 (0.05)**	0.20 (0.05)**	0.25 (0.06)**	1.16 (2; 0.56)
Social-emotional	0.25 (0.06)**	0.21 (0.05)**	0.19 (0.07)**	0.45 (2; 0.80)
Motivation	0.11 (0.07) [b]	0.05 (0.87) [b]	0.07 (0.09) [b]	0.14 (2; 0.93)
Other	0.39 (0.10)** [b]	0.24 (0.23) [b]	0.24 (0.16) [b]	0.87 (2; 0.65)

**p < 0.01; * *p* < 0.05; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions.

Table 9 *Average effects (SE) for socioeconomic status*

Outcome	Low SES	Mid and high SES	Meta-ANOVA Q-between (*df; p*)
Overall	0.20 ((0.03)**	0.21 (0.04)**	0.02 (1; 0.88)
Academic	0.15 (0.05)**	0.18 (0.04)** [a]	0.34 (1; 0.56)
Behaviour	0.21 (0.04)**	0.24 (0.04)**	0.23 (1; 0.63)
Social-emotional	0.21 (0.04)**	0.14 (0.07)*	0.95 (1; 0.33)
Motivation	0.13 (0.05)** [b]	-0.02 (0.07) [b]	3.35 (1; 0.07)
Other	0.38 (0.08)** [a]	0.07 (0.06) [b]	10.23 (1; 0.00)**

**p < 0.01; * *p* < 0.05; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions. Interventions with missing data on SES were excluded from the analysis.

Table 10 *Average effects (SE) for student behaviour*

Outcome	Regular	Behaviour problems	Meta-ANOVA
			Q-between (df; p)
Overall	0.20 (0.02)**	0.27 (0.08)**	0.62 (1; 0.43)
Academic	0.16 (0.04)**	0.50 (0.44) [b]	0.59 (1; 0.44)
Behaviour	0.22 (0.02)**	0.29 (0.08)**	0.72 (1; 0.40)
Social-emotional	0.19 (0.03)**	-	-
Motivation	0.08 (0.08)	-	-
Other	0.26 (0.10)*	-	-

**$p < 0.01$; * $p < 0.05$; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions. Interventions with missing data on student behaviour were excluded from the analysis.

Table 11 *Average effects (SE) for country*

Outcome	USA	Other country	Meta-ANOVA
			Q-between (df; p)
Overall	0.20 (0.03)**	0.26 (0.05)**	1.31 (1; 0.25)
Academic	0.16 (0.04)**	0.18 (0.07)** [b]	0.03 (1; 0.85)
Behaviour	0.21 (0.03)**	0.28 (0.05)**	1.24 (1; 0.27)
Social-emotional	0.20 (0.04)**	0.25 (0.07)**	0.40 (1; 0.53)
Motivation	0.08 (0.08)	-	-
Other	0.38 (0.08)** [a]	0.07 (0.06) [b]	10.23 (1; 0.00)**

**$p < 0.01$; * $p < 0.05$; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions.

We also investigated whether the intervention effect was related to the duration of the program. The results of this analysis are displayed in Table 12. Again, we found hardly any differences between the moderator variable and the intervention effects. We only found a small difference for the social-emotional outcomes. This significant difference is caused by the fact that short-term programs are very effective in enhancing social-emotional outcomes ($g = 0.83$). Again, however, this effect size was computed on the basis of a very small number of interventions.

Table 12 *Average effects (SE) for duration*

Outcome	< 13 weeks	13 weeks − 1 year	> 1 year	Meta-ANOVA Q-between (*df; p*)
Overall	0.19 (0.24)	0.23 (0.03)**	0.21 (0.03)**	0.18 (2; 0.92)
Academic	0.30 (0.13)* [b]	0.16 (0.05)**	0.16 (0.06)**	1.10 (2; 0.58)
Behaviour	0.29 (0.21) [a]	0.21 (0.03)**	0.25 (0.05)**	0.56 (2; 0.76)
Social-emotional	0.83 (0.22)** [b]	0.21 (0.04)**	0.19 (0.06)**	8.44 (2; 0.02)*
Motivation	0.05 (0.87) [b]	0.11 (0.07) [b]	0.07 (0.09) [b]	0.14 (2; 0.93)
Other	-	0.23 (0.12) [a]	0.37 (0.19) [b]	0.38 (1; 0.54)

$**p < 0.01$; $* p < 0.05$; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions.

The estimated intervention effect might relate to how the effect was measured. In many intervention studies, the effect was estimated using ratings by the teacher, the student, or an observer. Table 13 shows the average effect sizes that were found for each rater. We found significant differences between the raters for all outcomes together, and for behavioural outcomes. Students reported less improvement after following the program in comparison with reports filled in by teachers and observers.

Table 13 *Average effects (SE) for rater*

Outcome	Teacher	Student	Observer	Meta-ANOVA Q-between (*df; p*)
Overall	0.24 (0.03)**	0.16 (0.03)**	0.30 (0.07)**	6.27 (2; 0.04)*
Academic	0.12 (0.03)**	0.16 (0.04)**	-	0.37 (1; 0.54)
Behaviour	0.26 (0.03)**	0.10 (0.03)**	0.26 (0.07)**	13.44 (2; 0.00)**
Social-emotional	0.24 (0.04)**	0.18 (0.05)**	0.47 (0.49) [b]	1.06 (2; 0.60)
Motivation	0.09 (0.04)* [a]	0.10 (0.38) [a]	-	0.00 (1; 0.99)
Other	0.16 (0.06)** [b]	0.18 (0.16) [a]	0.39 (0.10)**	4.38 (2; 0.11)

$**p < 0.01$; $* p < 0.05$; [a] statistic in cell is based on no more than three or four interventions; [b] statistic in cell is based on only one or two interventions.

5   Discussion

5.1 Summary of the results

The meta-analysis included 54 classroom management interventions (presented in 47 different studies) aimed at enhancing students' academic, behavioural, social-emotional, motivational, or other related student outcomes. A large variety of interventions was implemented in the studies that met our inclusion criteria. Our analyses included five classroom management interventions that were implemented in at least three studies, namely, SWPBS, PATHS, GBG, Second Step, and Zippy's Friends. Yet together, they represented only 43% of the overall sample of selected studies.

In 85% of the included studies interventions were used that (among other foci) focused on changing students' behaviour, and 74% at least partly focused on improving students' social-emotional development. Half of the included studies reported on interventions that (at least partly) focused on changing the teachers' behaviour (54%). Only two intervention studies were explicitly focused on improving teacher-student relationships (4%). The most common combinations of classroom management components were a combination of focusing on students' behaviour and students' social-emotional development (18 studies), and these two student components combined with a teacher focus (13 studies). This trend towards more student-centred approaches rather than teacher-centred approaches is in line with the general tendency in primary education towards student-centred learning environments.

Across all interventions, we calculated an overall effect of Hedges' $g = 0.22$ on the various student outcomes (0.17 if the publication bias is taken into account). There were no significant differences between the various groups of outcomes: academic, behavioural, social-emotional, motivational, and other (e.g., time-on-task, self-efficacy, peer acceptance). Thus, the results of the meta-analysis confirm the finding of generally positive effects of classroom management interventions on student outcomes in primary education. In prior meta-analyses (Durlak et al., 2011; Marzano et al., 2003; Oliver et al., 2011), the reported effects were generally similar in size (i.e., when the effect sizes measured at the classroom level, such as in Marzano et al., 2003, or measured at the school level, such as in Oliver et al., 2011, are recalculated). Durlak et al. (2011) found somewhat larger effects for social-emotional outcomes (0.57) than we found in our study (0.22). Our meta-analysis included the recent literature only (published between 2003 and 2013). It is, therefore, noteworthy that our overall finding that classroom management interventions are generally effective in enhancing student outcomes is in line with the findings of prior meta-analyses which were mostly based on earlier publications.

To determine to which components of the classroom management interventions their effectiveness can be attributed, we performed several moderator analyses. The results indicate

that interventions focused on the social-emotional development of the students were somewhat more effective than interventions without this component. In particular, the social-emotional outcomes benefitted from this component. Furthermore, the programs that were most effective in enhancing students' academic performance were those that had a strong focus on improving the teachers' classroom management skills. This also applied to the outcome category 'other' (e.g., time-on-task, self-efficacy). Interventions focusing on multiple components had somewhat higher effects on the academic and 'other' outcomes, but not on the other outcome types. Furthermore, the exact combination of components on which programs focused had no influence on the intervention effect. Finally, we examined the effectiveness of the five intervention programs that were most common in our meta-analysis. We found that all programs were equally effective, except for SWPBS, which was not found to have an effect on the outcome measure 'all outcome types together'.

Additional moderator analyses revealed no large differences in the reported effects with respect to sex, socioeconomic status (low versus mid or high), student behaviour (regular or students with behavioural problems), grade year (pre-K to 1, 2 to 6, or both), or country (USA versus non-USA), indicating that all students may benefit from a classroom management intervention.

5.2 Scientific contribution

The findings of the present meta-analysis contribute to the current body of knowledge on classroom management by bringing together a broad span of recently conducted intervention studies on classroom management. In the selected studies, appropriate research designs were used to investigate the effects of various CMS/CMP on a variety of student outcomes. Whereas most prior researchers included studies without control groups in their meta-analyses, our focus was solely on studies with a control group. Therefore, maturation effects on social-emotional development, behaviour, and achievement were controlled for in designs with a control group. Hence, we can be confident that the reported effects on student outcomes were caused by the interventions. Moreover, a range of different student outcomes were used: academic, behavioural, social-emotional, motivational, and other relevant student outcomes. The fact that many studies included multiple outcome measures enabled us to evaluate the effects of the interventions on (almost) all these outcomes.

Another relevant point is that the studies we included were published in the last decade, and thus in current educational settings. In some studies, the data used were collected several years earlier; however, in most studies the data were collected in relatively modern classrooms. Furthermore, we paid specific attention to classroom management programs that are commonly used in educational practice. General descriptions of five of these programs are included in our paper (SWPBS, PATHS, GBG, Second Step, and Zippy's Friends); for four of these, Dutch equivalents are available (all except Second Step). As yet, the effectiveness of

several of these programs has not been investigated intensively. Although only a small number of studies of these programs could be included in our analyses (a minimum of three studies per program), we found that all programs (except SWPBS for 'all outcome types together') positively enhanced student outcomes.

5.3 Practical implications

Classroom management aims to facilitate both academic and social-emotional learning (Everston & Weinstein, 2006). In our meta-analysis, the strongest effects were found for programs targeting social-emotional development, particularly on the social-emotional outcome measure. This is considered a promising finding given that in current society, social skills are important for success later in the school career and in the work force (Jennings & DiPrete, 2010; Lynch & Simpson, 2010; Rhoades, Warren, Domitrovich, & Greenberg, 2011). Jennings and DiPrete (2011), for example, found that social and behavioral skills have a positive effect on the growth of academic skills in the early elementary grades. For the Netherlands, we consider this result to be especially relevant given the recent educational trend toward more inclusive classrooms. From August 2014, Dutch schools are legally obliged to offer suitable education to all students, including students with special needs (e.g., students with learning and/or behavioural problems, but also with physical and/or sensory hindrance). Consequently, more and more students with special needs will be integrated in regular classrooms. A typical classroom will increasingly consist of heterogeneous student groups with wide-ranging academic needs (Bosker & Doolaard, 2009; Veenman, Lem, Roelofs, & Nijsen, 2003). The challenges of this new policy cannot be fully anticipated, yet it seems apparent that students – in order to learn, work, and play together – will benefit from programs that focus on the development of empathy and social skills.

A second finding of this meta-analysis was that the interventions which focused on changing teachers' classroom management (e.g., keeping order, introducing rules and procedures, disciplinary interventions) had a small effect on students' academic outcomes. Classroom management is considered a precondition for learning; effective teaching and learning cannot take place in poorly managed classrooms (Jones & Jones, 2012). One can explain these findings through improved time-on-task, improved instruction practices and increased opportunity-to-learn, but this hypothesized causal chain needs to be further explored in future research (for more recommendations for future research, see section 5.4). Time-on-task was one of the outcomes we classified in the category 'other outcome measures'. Since this outcome measure was used in only a few studies, it was not feasible to analyse it separately. The category 'other outcome measures' also included outcome measures such as self-efficacy and peer acceptance. Given the importance of academic outcomes, as stressed by the Dutch Ministry, the Inspectorate of Education, and schools themselves (Ministry of Education, 2008; Inspectorate of Education, 2008, 2010; Vrielink, Hogeling, & Brukx, 2009),

more work is needed to understand how exactly student learning can be maximized through classroom management.

It must be remembered that most interventions (on average) showed positive effects on all student outcomes. Our findings clearly indicate that all students may benefit from these interventions. It is, however, essential that all stakeholders (policymakers, principals, teachers, and teacher educators) realize that the programs we investigated were often school-wide approaches in which a broad variety of strategies was used, indicating that there is no simple solution for classroom management problems.

5.4 Limitations and suggestions for further study

The studies included in the meta-analysis predominantly reported on the effectiveness of school-wide programs which had a broad focus on improving teaching practices, teacher-student relationships, student behaviour, and student social-emotional development. Although the effects of school-wide universal classroom management programs have often been investigated, few researchers have used pretest-posttest control group designs to estimate the effects on students' learning (both academic and social-emotional) and/or student behaviour (see also Chitiyo et al., 2012). Consequently, the number of studies with a broad focus that met our inclusion criteria was small, considering that 241 potential studies resulted from the literature search. Although the number of studies included was sufficient for the analyses, we would like to stress that the results should be interpreted with some caution. The findings showed that our meta-analysis was subject to some publication bias. Moreover, the findings of moderator analyses showed that students reported less enhancement by the interventions than was reported by teachers and observers, which might be caused by teachers' and observers' desire to find significant progress. Then again, self-reports of young students may be inaccurate if the research instruments are too complicated for them. Furthermore, we were unable to take all moderators into account in one single analysis, due to the relatively low number of studies that met the inclusion criteria for the meta-analysis.

With regard to the outcome measures, we would like to stress that various measures were used, for instance, for academic outcomes. The use of standardized tests was limited, which makes it difficult to generalize the results to all academic outcomes. Time-on-task, which we expected to be a relevant outcome measure, was not often measured. Furthermore, various instruments were used to measure student behaviour and students' social-emotional outcomes. Although we eliminated student outcomes measured using highly unreliable instruments, the construct validity of the various instruments was often unclear. As we have mentioned a number of times above, our results need to be interpreted with care.

Among the 241 potential studies, we found a wealth of small-scale studies in which multiple baseline designs were used to evaluate the effectiveness of a behavioural intervention (e.g., token economies, group contingencies, positive reinforcement) on student behaviour

among students with emotional and/or behavioural disorders (Cihak, Kirk, & Boon, 2009; Devender & Sokolosky, 2012; Fabiano & Pelham, 2003; Heering & Wilder, 2006; Kamps, Wendland, & Culpepper, 2006; Kamps et al., 2011; Kraemer, Davies, Arndt, & Hunley, 2012; Lambert, Cartledge, Heward, & Ya-yu, 2006; Lannie & McCurdy, 2007; Ling, Hawkins, & Weber, 2011; McGoey, Schneider, Rezzetano, Prodan, & Tankersley, 2010; Sutherland, Alder, & Gunter, 2003; Ya-yu & Cartledge, 2004), students with learning problems (e.g., Lohrmann & Talerico, 2004), or regular students, but generally without a control group (Carter & Van Norman, 2010; Christie & Schuster, 2003; Filcheck, McNeil, Greco, & Bernard, 2004; Franzen & Kamps, 2008; Little, Akin-Little, & Newman-Eig, 2010; McKissick, Hawkins, Lentz, Hailley, & McGuire, 2010; Stormont, Smith, & Lewis, 2007; Theodore, Dioguardi, Hughes, Aloiso, Carlo, & Eccles, 2009; Wright & McCurdy, 2012; Yarrough, Skinner, Lee, & Lemmons, 2004). Although these studies did not meet our inclusion criteria, their results, particularly those focusing on students with emotional and/or behavioural disorders, can be relevant to teachers facing the challenge of inclusive classrooms.[5]

Unfortunately, our literature search did not result in well-designed experimental studies in which the effectiveness of CMS/CMP was investigated in inclusive classrooms. Although such a study would clearly contribute to the field, it has been suggested that "teachers who believe students with special needs are their responsibility tend to be more effective overall with all of their students" (Jordan, Schwartz, & McGhie-Richmond, 2009, p. 535); this is in line with our finding that all students may benefit from a classroom management intervention. Nevertheless, further investigation of this topic is desirable. Although some attention has been given to the effects of the teacher on the functioning of students with special needs in mainstream classrooms (e.g., Wilkinson, 2005), researchers have generally not focused on the consequences of inclusive classrooms for teachers' required classroom management skills.

Another limitation of our meta-analysis is that we found only one study (Connor, Ponitz, Phillips, Travis, Glasney, & Morrison, 2010) that met our inclusion criteria and focused on the inclusion of ICT as (part of) a classroom management intervention. Little research exists that documents how the introduction of technology (e.g., interactive whiteboards, tablets) affects classroom management; there are few exceptions (e.g., Lim, Teo, Wong, Khine, Chai, & Divaharan, 2003). To our knowledge, no scientifically based research using pretest-posttest control group designs has been conducted to investigate how technology impacts on classroom management issues and practices. In 2006, Bolick and Cooper stated: "While adding technology to a classroom equips teachers with a new range of classroom management

---

[5] For recent reviews of studies using single case designs we refer to Solomon, Klein, Hintze, Cressey, and Peller (2012), Stage and Quiroz (1997), and Vannest, Davis, Davis, Mason, and Burke (2010). Solomon et al. (2012) examined the effects of school-wide positive behaviour support programs (SWPBS); Stage and Quiroz (1997) investigated the effects of disciplinary interventions on reducing the disruptive behaviour of individual students; and Vannest et al. (2010) investigated the effects of daily behaviour report cards (DBRC) on individual students' behaviour. Marzano et al.'s (2003) study also included single case studies.

tools such as spreadsheets and databases to manage school and classroom records and information, technology also presents a series of new classroom management issues such as moving students from the classroom to the computer lab or managing a classroom in which students are using a variety of different technologies such as wireless laptops or handheld computing devices" (pp. 541-542). As the aforementioned definition of classroom management pertains to the actions that teachers take to create an environment that facilitates both academic and social-emotional learning (Evertson & Weinstein, 2006), we see a promising role for technology-driven classroom management tools. Yet it remains unclear how they facilitate academic, behavioural, and social-emotional learning given the limited research in this area. One exception is the study by Ysseldyke, Spicuzza, Kosciolok, and Boys (2003), in which educational software (an instruction management system called "Accelerated Math") was used to increase students' academic learning time. The teachers used an automatized performance-tracking system that provided individualized practice for students on mathematics objectives, giving the teachers the opportunity to change their classroom organization to some extent (see also Burns, Klingbeil, & Ysseldyke, 2010; Nunnery & Ross, 2007; Ysseldyke, Spicuzza, Kosciolok, Teelucksingh, Boys, & Lemkuil, 2003; Ysseldyke & Tardrew, 2007 for similar quasi-experimental studies). Using computer-adaptive programs is, in our view, a promising innovation that could increase students' time-on-task because they do not need to wait for the teacher to provide them with practice worksheets that meet their academic needs. Another example is the study by Inan, Lowther, Ross, and Strahl (2010), which demonstrated that when computers are incorporated in the learning environment, classroom practices become more student-centred than teacher-centred (i.e., teachers as facilitators). However, it is evident that more research is needed to evaluate the impact of incorporating technology in primary school classrooms, and investigate the impact on various student outcomes.

It remains unclear whether the interventions we selected would have similar effects in societal contexts other than those in which they were implemented. In the Netherlands, for example, there are Dutch equivalents for a number of the programs we referred to in our study (SWPBS, GBG, PATHS, and Zippy's Friends). However, the effectiveness of these programs has hardly been investigated in Dutch schools. The only exceptions are the studies by Van Lier et al. (2005) and Witvliet et al. (2009), who investigated the effects of GBG, and the study by Louwe et al. (2007), who investigated the effects of PATHS; all of these were included in our analyses. Several other Dutch programs, which include relevant classroom management components, are used in primary schools; however, their effectiveness has not been scientifically examined, at least not using a suitable research design (e.g., with a control group). The program 'Leefstijl' is an example of such a Dutch program. 'Leefstijl' focuses on enhancing students' social-emotional competencies through a series of lessons and various group activities. The content of the program is in many ways similar to that of many of the

classroom management programs we found in the literature, yet, to our knowledge, its effectiveness has not been investigated in a scientific setting.

Another recommendation for further research pertains to the use of longitudinal studies. Out of the 241 potentially suitable publications, we found only two studies in which the long-term effects of a classroom management intervention (GBG) were measured ; that is, the effects of implementing the intervention in grades 1 and 2 on student outcomes during adolescence (Bradshaw, Zmuda, Kellam, & Lalongo, 2009; Kellam et al., 2008). More longitudinal studies are needed to investigate the maintenance effects of classroom management interventions, for example, by using follow-up tests on various student outcomes at different ages. Particularly the school-wide universal classroom management programs may have sustained effects on students' behaviour and social-emotional development, because these are relatively intensive programs.

Finally, we would like to present some recommendations for the scientific community on the basis of our experiences in reporting pretest-posttest control group designs used to evaluate the effectiveness of classroom management interventions. We found that numerous studies lacked detailed descriptions of the intervention that was implemented in the schools (e.g., specific focus of the teacher sessions and/or student sessions, type of training teachers and/or students received, duration of the intervention). Moreover, very few studies reported the classroom setting (e.g., group or frontal placement) in which the intervention was implemented, whereas such contextual factors may strongly influence student behaviour in the classroom. Similarly, it was often unclear within what type of school or educational context (e.g., during instruction, collaborative assignments, independent seatwork, or throughout the school day) the intervention was implemented. And when the intervention was implemented throughout the school day, it was unclear how the school days were normally organized (e.g., the amount of instruction time, independent seatwork, how often students worked collaboratively in groups, whether some students followed an individual learning trajectory, whether computers were used throughout the day, and whether teaching assistants were present). Information on these aspects makes the interpretation of the effectiveness of classroom management interventions much more insightful and, moreover, makes the findings much easier to replicate. We therefore strongly recommend including detailed descriptions of these aspects in scientific papers evaluating the effectiveness of CMS/CMP. Another recommendation is to provide detailed information on the research design and sampling procedures. On several occasions, it was unclear whether a control group was used, how the randomization or matching across intervention and control groups was performed, and whether the students were representative of the student population (e.g., many studies lacked details on gender, socioeconomic status, or ethnicity of the students included). In reporting the results, mean scores, standard deviations, and sample sizes among intervention and control groups should be reported for both pretest and posttest measures. Only then can effect sizes be

properly calculated. Moreover, for these measures, reliable and validated research instruments should be used (and information about this should be reported).

Despite the aforementioned limitations and the clear need for more high-quality program evaluations, sufficient evidence was found that several classroom management interventions lead to different types of outcomes for these interventions to be considered for implementation in primary school classrooms. As a result of this meta-analysis, preconditions for effective teaching and learning found in recent studies have been identified.

6   References

References marked with an asterisk indicate studies included in the meta-analysis.

Anderson, C. M., & Kincaid, D. (2005). Applying behavior analysis to school violence and discipline problems: Schoolwide positive behavior support. *The Behavior Analyst, 28*, 49-64.

*Arda, T. B., & Ocak, Ş. (2012). Social competence and promoting alternative thinking Strategies—PATHS preschool curriculum. *Educational Sciences: Theory & Practice, 12*, 2691-2698.

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behavior game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*, 119-124.

*Bavarian, N., Lewis, K. M., DuBois, D. L., Acock, A., Vuchinich, S., Silverthorn, N., Snyder, F. J., Day, J., & Ji, P. (2013). Using social-emotional and character development to improve academic outcomes: A matched-pair, cluster-randomized controlled trial in low-income, urban schools. *The Journal of School Health, 83*, 771-779.

*Benner, G. J., Nelson, J. R., Sanders, E. A., & Ralston, N. C. (2012). Behavior intervention for students with externalizing behavior problems: Primary-level standard protocol. *Exceptional Children, 78*, 181-198.

*Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2010). The effects of a multiyear universal social-emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology, 78,* 156-168.

*Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., Blair, C., Nelson, K. E., & Gill, S. (2008). Promoting academic and social-emotional school readiness: The head start REDI program. *Child Development, 79*, 1802-1817.

Bierman, K., Greenberg, M. T., & the Conduct Problems Prevention Research Group. (1996). Social skills training in the fast track program. preventing childhood disorders, substance abuse, and delinquency. In R. D. Peters, & R. J. McMahon (Eds.), *Preventing childhood disorders, substance abuse, and delinquency* (pp. 65-89). Thousand Oaks, CA: Sage.

Bohn, C. M., Roehrig, A. D., & Pressley, M. (2004). The first days of school in effective and less effective primary-grades classrooms. *Elementary School Journal, 104,* 269-287.

Bolick, C. M., & Cooper, J. M. (2006). Classroom management and technology. In: C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management. Research, practice, and contemporary issues* (pp. 541-558). New York / London: Lawrence Erlbaum Associates.

Borenstein, M., Hedges, L. V., Higgings, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis.* New York: Wiley.

Bosker, R. J. & Doolaard, S. (2009). De pedagogische kwaliteit van differentiatie in het onderwijs [The pedagogic quality of differentiation in education]. In A. Minnaert, H. Lutje Spelberg & H. Amsing (Eds.) *Het pedagogisch quotiënt. Pedagogische kwaliteit in opvoeding, hulpverlening, onderwijs en educatie* (pp. 151-168) . Houten, the Netherlands: Bohn Stafleu van Loghum.

Bradshaw, C. P. Zmuda, J. H., Kellam, S. G., & Lalongo, N. S. (2009). Longitudinal impact of two universal preventive interventions in first grade on educational outcomes in high school. *Journal of Educational Psychology, 101,* 926-937.

*Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. *Pediatrics, 130*, 1136-1145.

*Brigman, G. A., & Webb, L. D. (2003). Ready to learn: Teaching kindergarten students school success skills. *The Journal of Educational Research, 96*, 286-292.

Brophy, J. (2006). History of Research on Classroom Management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management. Research, practice, and contemporary issues* (pp.17-43). Malwah, NJ: Lawrence Erlbaum Associates.

Burns, M. K., Klingbeil, D. A., & Ysseldyke, J. (2010). The effects of technology-enhanced formative evaluation on student performance on state accountability math tests. *Psychology in the Schools, 47,* 582-591.

Carter, D. R., & Van Norman, R. K. (2010). Class-wide positive behavior support in preschool: Improving teacher implementation through consultation. *Early Childhood Education Journal, 38,* 279-288.

*Catalano, R. F., Mazza, J. J., Harachi, T. W., Abbott, R. D., Haggerty, K. P., & Fleming, C. B. (2003). Raising healthy children through enhancing social development in elementary school: Results after 1.5 years. *Journal of School Psychology, 41,* 143-164.

Chitiyo, M., May, M. E., & Chitiyo, G. (2012). An assessment of the evidence-base for school-wide positive behavior support. *Education and Treatment of Children, 35,* 1-24.

Christie, C.A., & Schuster, J.W. (2003). The effects of using response cards on student participation, academic achievement, and on-task behavior during whole-class, math instruction. *Journal of Behavioral Education, 12,*147-165.

Cihak, D. F., Kirk, E. R., & Boon, R. T. (2009). Effects of classwide positive peer "tootling" to reduce the disruptive classroom behaviors of elementary students with and without disabilities. *Journal of Behavioral Education, 18*(4), 267-278.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* New Jersey: Lawrence Erlbaum.

*Connor, C. M., Ponitz, C. C., Phillips, B. M., Travis, Q. M., Glasney, S., & Morrison, F. J. (2010). First graders' literacy and self-regulation gains: The effect of individualizing student instruction. *Journal of School Psychology, 48*, 433-455.

*Crean, H. F., & Johnson, D. B. (2013). Promoting alternative thinking strategies (PATHS) and elementary school aged children's aggression: Results from a cluster randomized trial. *American Journal of Community Psychology, 52,* 56-72.

*Curtis, C., & Norgate, R. (2007). An evaluation of the promoting alternative thinking strategies curriculum at key stage 1. *Educational Psychology in Practice, 23*, 33-44.

*Daunic, A., Smith, S., Brank, E., & Penfield, R. (2006). Classroom-based cognitive-behavioral intervention to prevent aggression: Efficacy and social validity. *Journal of School Psychology,44*, 123-139.

Devender, R. B., & Sokolosky, S. (2012). Effectiveness of noncontingent attention to decrease attention-maintained disruptive behaviors in the general education classroom. *Child & Family Behavior Therapy, 34,* 130-140.

Dignath, C., Buettner, G., & Langfeldt, H. (2008). How can primary school students learn self-regulated learning strategies most effectively?: A meta-analysis on self-regulation training programmes. *Educational Research Review*, *3*, 101-129.

*Dion, E., Roux, C., Landry, D., Fuchs, D., Wehby, J., & Dupéré, V. (2011). Improving attention and preventing reading difficulties among low-income first-graders: A randomized study. *Prevention Science, 12*, 70-79.

Dolan, L., Turkkan, J., Werthamer-Larsson, L., & Kellam, S. G. (1989). *The good behavior game training manual.* Baltimore, MD: Johns Hopkins Prevention Research Center.

*Domitrovich, C. E., Cortes, R. C., & Greenberg, M. T. (2007). Improving young children's social and emotional competence: A randomized trial of the preschool 'PATHS' curriculum. *The Journal of Primary Prevention, 28*, 67-91.

54

Dorst, A. van, Wiefferink, K., Dusseldorp, E., Galindo Garre, F., Crone, M., Paulussen, Th. (2008). *Preventie van pesten op basisscholen volgens de PRIMA-methode.* Leiden, the Netherlands: TNO.

Dufour, S., Denoncourt, J., & Mishara, B. L. (2011). Improving children's adaptation: New evidence regarding the effectiveness of zippy's friends, a school mental health promotion program. *Advances in School Mental Health Promotion, 4*, 18-28.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development, 82,* 405-432.

Elias, M. J., & Schwab, Y. (2006). From compliance to responsibility: Social and Emotional Learning and classroom management. In: C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management. Research, practice, and contemporary issues* (pp. 309-341). New York / London: Lawrence Erlbaum Associates.

Erikson, E. H. (1968). *Identity: youth and crisis.* New York: W. W. Norton.

Evertson, C. M. & Weinstein, C. S. (Eds.) (2006). *Handbook of classroom management. Research, practice, and contemporary issues.* Mahwah, NJ: Larence Erlbaum Associates, Inc.

Fabiano, G. A., & Pelham, W. E. J. (2003). Improving the effectiveness of behavioral classroom interventions for attention-deficit/hyperactivity disorder: A case study. *Journal of Emotional and Behavioral Disorders, 11*, 124-130.

Ferguson, C. J., Miguel, C. S., Kilburn, J. C., & Sanchez, P. (2007). The effectiveness of school-based anti-bullying programs: A meta-analytic review. *Criminal Justice Review, 32,* 401-414.

Filcheck, H., McNeil, C., Greco, L., & Bernard, R. (2004). Using a whole-class token economy and coaching of teacher skills in a preschool classroom to manage disruptive behavior. *Psychology in the Schools, 41*, 351-361.

*Flannery, D. J., Vazsonyi, A. T., Liau, A. K., Guo, S., Powell, K. E., Atha, H., Vesterdal, W., & Embry, D. (2003). Initial behavior outcomes for the PeaceBuilders universal school-based violence prevention program.*Developmental Psychology, 39*, 292-308.

Folmer-Annevelink, E., Doolaard, S., Mascareno, M., & Bosker, R. J. (2010). Class size effects on the number and types of student-teacher interactions in primary education. *Journal of Classroom Interactions, 45,* 30-38.

Franzen, K., & Kamps, D. (2008). The utilization and effects of positive behavior support strategies on an urban school playground. *Journal of Positive Behavior Interventions, 10*, 150-161.

Freiberg, H. J., & Lapointe, J. M. (2006). Research-based programs for preventing and solving discipline problems. In: C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management. Research, practice, and contemporary issues* (pp. 735-786). New York / London: Lawrence Erlbaum Associates.

*Frey, K. S., Nolen, S. B., Van Schoiack Edstrom, L., & Hirschstein, M. K. (2005). Effects of a school-based social-emotional competence program: Linking children's goals, attributions, and behavior. *Journal of Applied Developmental Psychology, 26,* 171-200.

Froyen, L. A., & Iverson, A. M. (1999). *Schoolwide and classroom management: The reflective educator-leader.* Upper Saddle River, NJ: Prentice-Hall.

Greenberg, M. T., & Kusché, C. A. (1993). *Promoting social and emotional development in deaf children: The PATHS project.* Seattle: University of Washington Press.

Greenberg, M. T., & Kusché, C. A. (2002). *Promoting alternative thinking strategies: Blueprint for violence prevention (2nd ed.).* Boulder: Institute of Behavioral Sciences, University of Colorado.

Greenberg, M. T., Bierman, K. L., Coie, J. D., Dodge, K. A., Lochman, J. E., McMahon, R. J., et al. (2010). The effects of a multiyear universal social-emotional learning program: The role of student and school characteristics. *Journal of Consulting and Clinical Psychology, 78*, 156-168.

Grossman, D. C., Neckerman, H. J., Koepsell, T. D., Liu, P. Y., Asher, K. N., Beland, K., Frey, K., & Rivara, F. P. (1997). Effectiveness of a violence prevention curriculum among children in elementary school: A randomized controlled trial. *JAMA, 277,* 1605-1611.

*Hamre, B. K., Pinata, R. C., Mashburn, A. J., & Downer, J. T. (2012). Promoting young children's social competence through the preschool PATHS curriculum and MyTeachingPartner professional development resources. *Early Education and Development, 23,* 809-832.

*Han, S. S., Catron, T., Weiss, B., Marciel, K. K. (2005). A teacher-consultation approach to social skills training for pre-kindergarten children: treatment model and short-term outcome effects. *Journal of Abnormal Child Psychology, 33,* 681-693.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement.* London; New York: Routledge.

Heering, P. W., & Wilder, D. A. (2006). The use of dependent group contingencies to increase on-task behavior in two general education classrooms. *Education and Treatment of Children, 29,* 459-467.

*Holen, S., Waaktaar, T., Lervåg, A., & Ystgaard, M. (2012). The effectiveness of a universal school-based programme on coping and mental health: a randomised, controlled study of Zippy's Friends. *Educational Psychology, 32,* 657-677.

*Holen, S., Waaktaar, T., Lervåg, A., & Ystgaard, M. (2013). Implementing a universal stress management program for young school children: Are there classroom climate or academic effects? *Scandinavian Journal of Educational Research, 57,* 420-444.

*Holsen, I., Smith, B. H., & Frey, K. S. (2008). Outcomes of the social competence program Second Step in Norwegian elementary schools. *School Psychology International, 29,* 71-88.

*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,* 133-144.

Inan, F. A., Lowther, D. L., Ross, S. M., & Strahl, D. (2010). Pattern of classroom activities during students' use of computers: Relations between instructional strategies and computer applications. *Teaching and Teacher Education, 26,* 540-546.

Inspectorate of Education. (2008). *De staat van het onderwijs. jaarverslag 2006-2007* [The state of education. Year report 2006-2007]. Utrecht, the Netherlands: Onderwijsinspectie.

Inspectorate of Education. (2010). *Opbrengstgericht werken in het basisonderwijs* [Data-driven teaching in elementary schools]. Utrecht, the Netherlands: Onderwijsinspectie.

Jennings, J. L., & DiPrete, T. A. (2010). Teacher effects on social and behavioral skills in early elementary school. *Sociology of Education, 83,* 135-159.

*Jones, S. M., Brown, J. L., & Lawrence Aber, J. (2011). Two-year impacts of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development, 82,* 533-554.

Jones, V.F. & Jones, L. S. (2012). *Comprehensive classroom management, creating communities of support and solving problems (10th ed).* Upper Saddle River, NJ: Pearson.

Jordan, A., Schwartz, E., & McGhie-Richmond, D. (2009). Preparing teachers for inclusive classrooms. *Teaching and Teacher Education, 25,* 535-542.

Kagan, S. (2005). *Structureel coöperatief leren* [Structural cooperative learning]. Vlissingen, the Netherlands: Bazalt.

Kamps, D., Wendland, M., & Culpepper, M. (2006). Active teacher participation in functional behavior assessment for students with emotional and behavioral disorder risks in general education classrooms. *Behavioral Disorders, 31,* 128-146.

Kamps, D., Wills, H. P., Heitzman-Powell, L., Laylin, J., Szoke, C., Petrillo, T., & Culey, A. (2011). Class-wide function-related intervention teams: Effects of group contingency programs in urban classrooms. *Journal of Positive Behavior Interventions, 13,* 154-167.

Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A large-scale evaluation of the KiVa anti-bullying program. *Child Development, 82,* 311-330.

Kellam, S. G., Brown, C. H., Poduska, J., Lalongo, N., Wang, W., Toyinbo, P., Petras, H., Ford, C., Windham, A., & Wilcox, H. C. (2008). Effects of a universal classroom behaviour management program in first and second grades on young adult behavioural, psychiatric, and social outcomes. *Drug Alcohol Dependence, 95,* S1-S28.

Kelshaw-Levering, K., Sterling-Turner, H. E., & Henry, J. R. (2000). Randomized interdependent group contingencies: Group reinforcement with a twist. *Psychology in the Schools, 37,* 523-533.

Klamer-Hoogma, M. (2012). *Klassenmanagement* [Classroom management]. Groningen/Houten, the Netherlands: Noordhoff Uitgevers.

*Koglin, U., & Petermann, F. (2011). The effectiveness of the behavioural training for preschool children. *European Early Childhood Education Research Journal, 19,* 97-111.

Kraemer, E. E., Davies, S. C., Arndt, K. J., & Hunley, S. (2012). A comparison of the mystery motivator and the *Get'em on task* interventions for off-task behaviors. *Psychology in the Schools, 49,* 163-175.

*Kutnick, P., Ota, C., & Berdondini, L. (2008). Improving the effects of group working in classrooms with young school-aged children: facilitating attainment, interaction and classroom activity. *Learning and Instruction, 18,* 83-95.

Lambert, M. C., Cartledge, G., Heward, W. L., & Ya-yu, L. (2006). Effects of response cards on disruptive behavior and academic responding during math lessons by fourth-grade urban students. *Journal of Positive Behavior Interventions, 8,* 88-99.

Lane, K. L. & Kuiper, I. (2012). *Voorkom probleemgedrag: strategieën die werken voor de basisschool.* Amsterdam, SWP.

Lane, K. L., Menzies, H. M., Bruhn, A. L., & Crnobori, M. (2011). *Managing challenging behaviors in schools. Research-based strategies that work.* New York: The Guilford Press.

Lannie, A. L., & McCurdy, B. L. (2007). Preventing disruptive behavior in the urban classroom: Effects of the Good Behavior Game on student and teacher behavior. *Education and Treatment of Children, 30,* 85-98.

Leflot, G., Van Lier, P. A. C., Onghena, P., & Colpin, H. (2013). The role of children's on-task behavior in the prevention of aggressive behavior development and peer rejection: A randomized controlled study of the good behavior game in Belgian elementary classrooms. *Journal of School Psychology, 51*, 187-199.

*Leflot, G., van Lier, Pol A. C., Onghena, P., & Colpin, H. (2010). The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the good behavior game. *Journal of Abnormal Child Psychology, 38*, 869-882.

Lewis, T. J. & Sugai, G. (1999). Effective behavior support: A systems approach to proactive schoolwide management. *Focus on Exceptional Children, 31,* 1-24.

Lim, C. P., Teo, Y. H., Wong, Ph., Khine, M. S., Chai, C. S., & Divaharan, S. (2003). Creating a conducive learning environment for the effective integration of ICT: Classroom management issues. *Journal of Interactive Learning Research, 14,* 405-423.

*Linares, L. O., Rosbruch, N., Stern, M. B., Edwards, M. E., Walker, G., Abikoff, H. B., Alvir, J., & Ma. J. (2005). Developing cognitive-social-emotional competencies to enhance academic learning. *Psychology in the Schools, 42,* 405-417.

Ling, S., Hawkins, R. O., & Weber, D. (2011). Effects of a classwide interdependent group contingency designed to improve the behaviour of an at-risk student. *Journal of Behavioral Education, 20,* 103-116.

Little, S. G., Akin-Little, A., & Newman-Eig, L. M. (2010). Effects on homework completion and accuracy of varied and constant reinforcement within an interdependent group contingency system. *Journal of Applied School Psychology, 26,* 115-131.

Lohrmann, S., & Talerico, J. (2004). Anchor the boat. A classwide intervention to reduce problem behavior. *Journal of Positive Behavior Interventions, 6,* 113-120.

*Louwe, J. J., Van Overveld, C. W., Merk, W., De Castro, B. O., & Koops, W. (2007). De invloed van het programma alternatieve denkstrategieën op reactieve en proactieve agressie bij jongens in het primair onderwijs: Effecten na één jaar [The effects of the PATHS-curriculum on reactive and proactive aggression]. *Pedagogische Studiën, 84*, 277-292.

*Lynch, K. B., Geller, S. R., Schmidt, M. G. (2004). Multi-year evaluation of the effectiveness of a resilience-based prevention program for young children. *The Journal of Primary Prevention, 24,* 335-353.

Lynch, S. A., & Simpson, C. G. (2010). Social skills: Laying the foundation for success. *Dimensions of Early Childhood, 38 (2),* 3-12.

Marzano, R. J. (2003). *What works in schools. Translating research into action.* Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

Marzano, R. J., Marzano, J. S., & Pickering, D. J. (2003). *Classroom management that works. Research-based strategies for every teacher.* Alexandria, VA: Association for Supervision and Curriculum Development (ASCD).

*McClowry, S. G., Snow, D. L., Tamis-LeMonda, C., & Rodriguez, E. T. (2010). Testing the efficacy of INSIGHTS on student disruptive behavior, classroom management, and student competence in inner city primary grades. *School Mental Health, 2,* 23-35.

McCurdy, B. L., Mannella, M. C., & Eldridge, N. (2003). Positive behavior support in urban schools: Can we prevent the escalation of antisocial behavior? *Journal of Positive Behavior Interventions, 5*, 158-170.

McGoey, K. E., Schneider, D. L., Rezzetano, K. M., Prodan, T., & Tankersley, M. (2010). Classwide intervention to manage disruptive behavior in the Kindergarten classroom. *Journal of Applied School Psychology, 26,* 247-261.

McKissick, C., Hawkins, R. O., Lentz, F. E., Hailley, J., & McGuire, S. (2010). Randomizing multiple contingency components to decrease disruptive behaviors and increase student engagement in an urban second-grade classroom. *Psychology in the Schools, 47,* 944-959.

McMahon, S. D., Washburn, J., Felix, E. D., Yakin, J., & Childrey, G. (2000). Violence prevention: Program effects on urban preschool and kindergarten children. *Applied and Preventive Psychology, 9*, 271-281.

Merrell, K. W., Gueldner, B. A., Ross, S. W., Isava, D. M. (2008). How effective are school bullying intervention programs? A meta-analysis of intervention research. *School Psychology Quarterly, 23,* 26-42.

Ministry of Education. (2008). *Beleidsreactie doorlopende leerlijnen taal en rekenen* [Policy response continuous learning trajectories language arts and mathematics]. Den Haag, the Netherlands: Ministerie van OCW.

*Mishara, B., & Ystgaard, M. (2006). Effectiveness of a mental health promotion program to improve coping skills in young children: Zippy's Friends. *Early Childhood Research Quarterly, 21,* 110-123.

Monkeviciené, O., Mishara, B. L., & Dufour, S. (2006). Effects of the Zippy's friends programme on childrens coping abilities during the transition from kindergarten to elementary school. *Early Childhood Education Journal, 34,* 53-60.

*Morris, P., Millenky, M., Raver, C. C., & Jones, S. M. (2013). Does a preschool social and emotional learning intervention pay off for classroom instruction and children's behavior and academic skills? Evidence from the foundations of learning project. *Early Education and Development, 24,* 1020-1042.

Nationaal Jeugdinstituut. (2013). *Naar meer wenselijk gedrag op de basisschool. Wat werkt?* [To more desirable behaviour in primary education. What works?]. Utrecht, the Netherlands: Nederlands Jeugdinstituut.

Nunnery, J. A., & Ross, S. M. (2007). The effects of the School Renaissance Program on student achievement in reading and mathematics. *Research in the Schools, 14,* 40-59.

O'Neill, S., & Stephenson, J. (2012). Does classroom management coursework influence pre-service teachers' perceived preparedness or confidence? *Teaching and Teacher Education, 28,* 1131-1143.

Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011). Teacher classroom management practices: effects on disruptive or aggressive student behavior. *Campbell Systematic Reviews, 4,* 1-55.

*Petermann, F., & Natzke, H. (2008). Preliminary Results of a Comprehensive Approach to Prevent Antisocial Behaviour in Preschool and Primary School Pupils in Luxembourg. *School Psychology International, 29,* 606-626.

Peters, J. H. (2012). Are they ready? Final year pre-service teachers' learning about managing student behaviour. *Australian Journal of Teacher Education, 37,* 18-42.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61,* 991-996.

*Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Metzger, M. W., & Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: a cluster-randomized control trial. *Journal of Consulting and Clinical Psychology, 77,* 302-316.

Rhoades, B. L., Warren, H. K., Domitrovich, C. E., & Greenberg, M. T. (2011). Examining the link between preschool social-emotional competence and first grade academic achievement: The role of attention skills. *Early Childhood Research Quarterly, 26,* 182-191.

*Riggs, N. R., Greenberg, M. T., Kusché, C. A., & Pentz, M. A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: Effects of the PATHS curriculum. *Prevention Science, 7*, 91-102.

*Rimm-Kaufman, S. E., Fan, X., Chiu, Y.-J., & You, W. (2007). The contribution of the Responsive Classroom Approach on children's academic achievement: Results from a three year longitudinal study. *Journal of School Psychology, 45,* 401-421.

Rydell, A. M., & Henricsson, L. (2004). Elementary school teachers' strategies to handle externalizing classroom behaviour: a study of relations between perceived control, teacher orientation and strategy preferences. *Scandinavian Journal of Psychology, 45,* 93-102.

*Schick, A., & Cierpka, M. (2005). Faustlos: Evaluation of a curriculum to prevent violence in elementary schools. *Applied and Preventive Psychology, 11,* 157-165.

Schussler, D. L., Poole, I. R., Whitlock, T. W., & Evertson, C. M. (2007). Layers and links: Learning to juggle "one more thing" in the classroom. *Teaching and Teacher Education, 23,* 572-585.

*Seifer, R., Gouley, K., Miller, A. L., & Zakriski, A. (2004). Implementation of the PATHS curriculum in an urban elementary school. *Early Education and Development, 15*, 471-485.

*Self-Brown, S., & Mathews, S. (2003). Effects of classroom structure on student achievement goal orientation. *Journal of Educational Research, 97*, 106-111.

Shook, A. C. (2012). A study of preservice educators' dispositions to change behavior management strategies. *Preventing School Failure: Alternative Education for Children and Youth, 56,* 129-136.

Smart, J. B., & Brent, I. L. (2010). A grounded theory of behavior management strategy selection, implementation, and perceived effectiveness reported by first-year elementary teachers. *Elementary School Journal, 110,* 567-584.

*Snyder, J., Low, S., Schultz, T., Barner, S., Moreno, D., Garst, M., Leiker, R., Swink, N., & Schrepferman, L. (2011). The impact of brief teacher training on classroom management and child behavior in at-risk preschool settings: Mediators and treatment utility. *Journal of Applied Developmental Psychology, 32*, 336-346.

Solomon, B. G., Klein, S. A., Hintze, J. M., Cressey, J. M., & Peller, S. L. (2012). A meta-analysis of schoolwide positive behavior support: An exploratory study using single case synthesis. *Psychology in the Schools, 49,* 105-122.

*Sørlie, M., & Ogden, T. (2007). Immediate impacts of PALS: A school-wide multi-level programme targeting behaviour problems in elementary school. *Scandinavian Journal of Educational Research, 51*, 471-492.

Spilt, J. L., Koot, H. M., & Van Lier, P. A. C. (2013a). Wie profiteert van taakspel? [Whom benefits from 'Taakspel']. *Kind & Adolescent, 34*, 195-207.

Spilt, J. L., Koot, J. M., & Van Lier, P. A. C. (2013b). For whom does it work? Subgroup differences in the effects of a school-based universal prevention program. *Prevention Science, 14*, 479-488.

Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review, 26,* 333-368.

Stormont, M. A., Smith, S. C., & Lewis, T. J. (2007). Teacher implementation of precorrection and praise statements in head start classrooms as a component of a program-wide system of positive behavior support. *Journal of Behavioral Education, 16*, 280-290.

Sucuogly, B., Akalin, S., & Sazak-Pinar, E. (2010). The effects of classroom management on the behaviors of students with disabilities in inclusive classrooms in Turkey. *Journal of the International Association of Special Education, 11,* 64-74.

Sugai, G., & Horner, R. R. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child and Family Behavior Therapy, 24*, 23-50.

Sugai, G., & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review, 35*, 245-259.

Sutherland, K. S., Alder, N., & Gunter, R L. (2003). The effect of varying rates of opportunities to respond to academic requests on the behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 11,* 239-248

Swinson, J., Woof, C., & Melling, R. (2003). Including emotional and behavioural difficulties pupils in a mainstrain comprehensive: A study of the behaviour of pupils and classes. *Educational Psychology in Practice, 19,* 65-75.

Teitler, P., & Brussel, A. van. (2012). *Lessen in orde op de basisschool. Handboek voor de onderwijspraktijk* [Lessons in control in primary education]. Bussum, the Netherlands: Coutinho bv.

Theodore, L. A., DioGuardi, R. J., Hughes, T. L., Aloiso, D., Carlo, M., & Eccles, D. (2009). A class-wide intervention for improving homework performance. *Journal of Educational & Psychological Consultation, 19*, 275-299.

*Tiano, J. D., & McNeil, C. B. (2006). Training head start teachers in behavior management using parent-child interaction therapy: A preliminary investigation. *Journal of Early and Intensive Behavior Intervention, 3*, 220-233.

Ttofi, M. M. & Farrington, D. P. (2011). Effectiveness of school-based programs to reduce bullying: a systematic and meta-analytic review. *Journal of Experimental Criminology, 7,* 27-56.

Van de Grift, W., Van der Wal, M., & Torenbeek, M. (2011). Ontwikkeling in de pedagogisch didactische vaardigheid van leraren in het basisonderwijs [Development in teaching skills]. *Pedagogische Studiën, 88,* 416-432.

Van der Sar, R. M. (2004). *Met taakspel lukt het wel. Een tussenrapportage over de effecten van taakspel op taakgericht en regelovertredend gedrag in de klas* [With 'taakspel' it is accomplishable]. Rotterdam, the Netherlands: Pedologisch Instituut, CED-Groep.

*Van Lier, P. A. C., Vuijk, P., & Crijnen, A. A. M. (2005). Understanding mechanisms of change in the development of antisocial behavior: The impact of a universal intervention. *Journal of Abnormal Child Psychology, 33,* 521-535.

Vannest, K. J., Davis, J. L., Davis, C. R., Mason, B. A., & Burke, M. D. (2010). Effective intervention for behavior with a daily behavior report card: A meta-analysis. *School Psychology Review, 39,* 654-672.

Veenman, S., Lem, P., Roelofs, E., & Nijsen, F. (2003). *Effectieve instructie en doelmatig klassemanagement. Een schoolverbeteringsprogramma voor enkelvoudige en combinatieklassen* [Effective instruction and purposive classroom management. A school improvement program for single-grade and multi-grade classrooms]. Lisse: Swets & Zeitlinger.

*Vespo, J. E., Capece, D., & Behforooz, B. (2006). Effects of the nurturing curriculum on social, emotional, and academic behaviors in kindergarten classrooms. *Journal of Research in Childhood Education, 20*, 275-285.

Vrielink, S., Hogeling, L. & Brukx, D. (2009) *Opiniepeiling Kwaliteitsagenda PO Eindrapportage.* Nijmegen, the Netherlands: ResearchNed.

*Webster-Stratton, C., Reid, M. J., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: Evaluation of the incredible years teacher and child training programs in high-risk schools. *Journal of Child Psychology and Psychiatry, 49*, 471-488.

Weissberg, R. P., Resnik, H., Payton, J., & O'Brien, M. U. (2003). Evaluating social and emotional learning programs. *Educational Leadership, 60,* 46-50.

Wilkinson, L. A. (2005). An evaluation of conjoint behavioral consultation as a model for supporting students with emotional and behavioral difficulties in mainstream classrooms. *Emotional and Behavioural Difficulties, 10,* 119-136.

Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine, 33,* 130-143.

Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive and disruptive behavior: a meta-analysis. *Journal of Consulting and Clinical Psychology, 71,* 136-149.

*Witvliet, M., van Lier, P. A. C., Cuijpers, P., & Koot, H. M. (2009). Testing links between childhood positive peer relations and externalizing outcomes through a randomized controlled intervention study. *Journal of Consulting and Clinical Psychology, 77*, 905-915.

Witvliet, M., Van Lier, P., Cuijpers, P., & Koot, H. (2010). Het stimuleren van contact met klasgenoten [Stimulating contact with classmates]. *Kind & Adolescent Praktijk, 9*, 168-173.

Woodcock, S. & Reupert, A. (2012). A cross-sectional study of student teachers' behaviour management strategies throughout their training years. *Australian Educational Researcher, 39,* 159-172.

Wright, R. A., & McCurdy, B. L. (2012). Class-wide positive behavior support and group contingencies: Examining a positive variation of the good behavior game. *Journal of Positive Behavior Interventions, 14*, 173-180.

Wubbels, T., Brekelmans, M., Van Tartwijk, J., & Admiraal, W. (1999). Interpersonal relationships between teachers and students in the classroom. In H.C. Waxman & H.J. Walberg (Eds.), *New directions for teaching practice and research* (pp. 151-170). Berkeley, CA: McCutchan.

Wubbels, T., Den Brok, P., Veldman, I., & Van Tartwijk, J. (2006). Leraar-leerlingrelaties en klassenmanagement in het hedendaagse onderwijs [Teacher-student relationships and classroom management in present-day education]. In R. Klarus & W. Wardekker (Eds.), *Wat is goed onderwijs: Bijdragen vanuit de pedagogiek* (pp. 121-141). Den Haag: Boom Lemma Uitgevers.

Yarrough, J. L., Skinner, C. H., Lee, Y. J., & Lemmons, C. (2004). Decreasing transition times in a second grade classroom: Scientific support for the timely transitions game. *Journal of Applied School Psychology, 20*(2), 85-107.

Ya-yu, L., & Cartledge, G. (2004). Total class peer tutoring and interdependent group oriented contingency: Improving the academic and task related behaviors of fourth-grade urban students. *Education and Treatment of Children, 27,* 235-262.

Ysseldyke, J., & Tardrew, S. (2007) Use of a progress monitoring system to enable teachers to differentiate mathematics instruction. *Journal of Applied School Psychology, 24,* 1-28.

Ysseldyke, J., Spicuzza, R., Kosciolok, S., & Boys, C. (2003). Effects of a learning information system on mathematics achievement and classroom structure. *Journal of Educational Research, 96,* 163-173.

Ysseldyke, J., Spicuzza, R., Kosciolok, S., Teelucksingh, E., Boys, C., & Lemkuil, A. (2003). Using a curriculum-based instructional management system to enhance math achievement in urban schools. *Journal of Education for Students Placed at Risk, 8,* 247-265.