Educ. Treat. Child. (2024) 47:83–104
https://doi.org/10.1007/s43494-024-00122-3

REVIEW

Behavior Change Potential of Classroom Behavior Management Mobile Applications: A Systematic Review

Sharon R. Mittiga · Nerelie C. Freeman · Erin S. Leif · Brett E. Furlonger · Ellen Pattison

Accepted: 31 January 2024 / Published online: 26 February 2024

Abstract Advances in classroom behavior management mobile applications (CBM apps) have led some teachers to use them to supplement their existing student management strategies, although little is known about their effectiveness in facilitating behavior change. This systematic review aimed to identify and appraise research on the effectiveness of CBM apps for promoting positive behavioral and learning outcomes of elementary, middle, and high school students. A systematic search was conducted in the PsycINFO, ERIC, and EBSCOhost databases for articles published between 2007 and 2020. The 15 included studies provided some preliminary evidence for CBM apps. Apps were primarily used to facilitate the delivery of self-monitoring interventions (SCORE IT and I-Connect) or class-wide reinforcement systems (ClassDojo and the Classroom Behavior Management System). An evaluation of study quality using the What Works Clearinghouse design standards (version 5.0) yielded mixed results, with only 53% of the included studies meeting standards with or without reservations. In general, these studies showed limited risk of bias and moderate to strong effect sizes. Based on the findings of the review, we provide practice recommendations and describe areas for future research.

S. R. Mittiga (✉) · N. C. Freeman · E. S. Leif · B. E. Furlonger · E. Pattison
Faculty of Education, Monash University, Melbourne, VIC, Australia
e-mail: sharon.mittiga@gmail.com

Keywords Classroom behavior management · Mobile applications · Student behavior · Behavior management strategies · Classroom interventions

Introduction

Classroom behavior management (CBM) has been a topic of significant interest among researchers for several decades (e.g., Evertson & Weinstein, 2006; Gettinger, 1988; Korpershoek et al., 2014) as teachers often face challenges in managing off-task student behavior. Classroom behavior management (CBM) refers to the techniques, strategies, and interventions employed by teachers to promote a positive learning environment by reducing or preventing disruptive behavior. These methods can include positive reinforcement, setting clear expectations, implementing consistent rules, and employing proactive measures (e.g., seating arrangements, hand signals, giving students choice, visual schedules) to engage students and reduce potential distractions (Evertson & Weinstein, 2006; Great Schools Partnership, 2014; Korpershoek et al., 2014). Studies have shown that disruptive behaviors, such as talking out of turn or being out of one's seat, are prevalent (Beaman et al., 2007; Clunies-Ross et al., 2008; Kaufman et al., 2010; Sullivan et al., 2014; Walter et al., 2006) and can disrupt the learning climate in the classroom. This, in turn, may negatively affect academic achievement (Alatupa et al., 2011; Greenwood et al., 1984; Marzano & Marzano, 2003; Shinn et al., 1987; Thomas et al., 2012). Studies have revealed that teachers who spend more time managing off-task or disruptive behavior have less time for academic instruction and may experience elevated stress levels (Beaman et al., 2007; Clunies-Ross et al., 2008; Ratcliff et al., 2010). Furthermore, a review by Ingersoll and Smith (2003) found that early career teachers' competency in CBM significantly influences their retention in the profession. However, teachers consistently report feeling underprepared to manage disruptive classroom behavior, citing lack of time, difficulty in implementing behavior management systems, and inadequate training as significant challenges (Fox et al., 2021).

Classroom Behavior Management (CBM) Mobile Applications

The utilization of technology to support the delivery of CBM has garnered growing attention in recent years, as evidenced by the proliferation of technology-based programs, such as mobile applications (apps), aimed at assisting teachers with promoting academically engaged behaviors (Hammonds et al., 2013; Riden et al., 2019; Robacker et al., 2016). For the purpose of the present review, CBM apps are defined as technology-based programs that utilize apps to support teachers with promoting positive student behaviors (Riden et al., 2019). Despite limited research on the efficacy of CBM apps, educational researchers have recommended apps such as ClassDojo and SCORE IT to support teachers with classroom management (Hammonds et al., 2013; O'Brien & Aguinaga, 2014; Riden et al., 2019). ClassDojo (ClassDojo Inc., 2020) is a CBM tool that enables teachers to set student goals, track behavior, and monitor student progress. It includes a customizable reward system that teachers can use to reinforce positive student behavior. ClassDojo allows teachers to provide visual and auditory feedback (e.g., rings, buzz sounds, Dojo points) to students. The app can also aid communication between teachers and parents regarding student behavior and progress. The self-monitoring app SCORE IT (Lizzy B. Good Behavior Consulting, 2015) enables teachers to promote positive behavior changes through antecedent-based strategies such as prompting, goal setting, and self-monitoring (Riden et al., 2019). Teachers can set individualized goals for students and prompt students to rate their adherence to behavior expectations on a sliding scale at predetermined intervals. Teachers can then provide feedback to students using the data collected via the app for correct self-monitoring, the occurrence of positive behavior, or for meeting their individual goals.

Classroom Behavior Management (CBM) Apps and Behavior Change Principles

A cursory review of CBM apps suggests that they have been developed with behavior change principles in mind (Riden et al., 2019; Robacker et al., 2016). These principles involve creating a positive and structured learning environment that promotes desirable classroom behavior (e.g., on-task, raising hand when seeking help) and discourages undesirable classroom behavior (e.g., off-task, out of seat; Lewis, 2008; Lewis et al., 2013). Antecedent-based strategies, such as clear expectations, cues, and self-monitoring, and consequence-based strategies, such as positive reinforcement and response cost, are commonly used (Murphy et al., 2019; Parsonson, 2012). For example, ClassDojo allows participants to exchange points for reinforcers (e.g., Dadakhodjaeva, 2017; Ford, 2018; Lipscomb et al., 2018) simulating a token economy (Riden et al., 2019; Robacker et al., 2016), which has been shown to be effective in promoting positive behavior (Pokorski et al., 2019; Robichaux & Gresham, 2014). However, to avoid an overreliance on extrinsic rewards, it may be beneficial to implement a gradual fading of rewards as the student achieves mastery of the task or behavior. A response cost system, a strategy that deducts points for undesirable classroom behavior, is sometimes used as part of a token economy but some researchers (e.g., DeJager et al., 2020) found it to be less effective when included in a token economy. ClassDojo has also been used in conjunction with a group contingency strategy, The Good Behavior Game (GBG), with studies showing that ClassDojo with GBG can be effective in reducing unwanted classroom behavior (Chaffee et al., 2017; Maggin et al., 2012; Riden et al., 2019; Simonsen et al., 2008). Other CBM apps have incorporated elements of self-management. For example, self-monitoring is a key characteristic of SCORE IT, in which students self-observe and record their behaviors when prompted by the app (Riden et al., 2019). Setting specific and challenging goals has also been shown to improve performance (Locke et al., 1981;

Lunenburg, 2011), and this is a strategy incorporated into some CBM apps. For example, when using ClassDojo and SCORE IT apps, teachers are able to set individual goals for students (Robacker et al., 2016).

Effectiveness of Classroom Behavior Management (CBM) Apps

Results from preliminary research on the effectiveness of CBM apps for facilitating student behavior change is mixed. Two studies that evaluated ClassDojo combined with the GBG reported an increase in students' academically engaged behaviors and a decrease in disruptive behaviors (Dadakhodjaeva, 2017; Ford, 2018). However, treatment integrity was observed to be low. For example, teachers in Dadakhodjaeva (2017) made mistakes when implementing the intervention, such as failing to announce the end of the GBG and incorrectly determining the winning team. A technical malfunction in the Ford (2018) study meant that the intervention was implemented *without* ClassDojo for an entire phase in two classrooms.

In another example, Lipscomb et al. (2018) evaluated ClassDojo and ClassDojo plus Tootling (in which students delivered Dojo points to their peers for appropriate behavior) and found reductions in disruptive behaviors displayed by seven university students with intellectual disabilities. It is interesting that Lipscomb et al. (2018) found that ClassDojo alone was more effective than ClassDojo plus Tootling. The authors noted that, for some participants, levels of disruptive behavior were low during baseline. In addition, the authors hypothesized that the peer-mediated intervention component (tootling) was less preferred by the students. When ClassDojo was implemented without verbal feedback or praise, both Elliott (2017) and Wilson (2017) found no significant changes in students' behaviors. Reasons cited for the null finding was that the primary school students already had a good relationship with their teachers (Elliott, 2017) and the kindergarten children were already well-behaved (Wilson, 2017). By contrast, Saeger (2017) evaluated ClassDojo plus response cost system but in the absence of back up reinforcers, and observed an increase in students' positive behaviors and a decrease in their inappropriate behaviors. Saeger noted that goal setting and self-assessment was also introduced during the intervention phase, allowing students to select a positive behavior they wished to work on and to determine if their goal was met at the end of each week. These additional intervention components may have enhanced the effectiveness of the intervention.

Research on the SCORE IT app has also produced mixed results. Vogelgesang et al. (2016) examined the effectiveness of the SCORE IT app in the absence of programmed consequences for accurate self-monitoring or appropriate behavior. They found that the introduction of SCORE IT was associated with an increase in academic engagement, with gains maintained at 2- and 4-week follow-ups. Other researchers have evaluated SCORE IT paired with antecedent-based strategies (prompting and self-monitoring) and consequence-based strategies (teacher praise, feedback and back-up reinforcers; e.g., Bruhn et al., 2016; Bruhn et al., 2015; Bruhn et al., 2017) and reported mixed findings. In Bruhn et al. (2015), the use of SCORE IT was associated with decreases in disruptive behavior for two participants but an increase in academic engagement for only one participant. One participant showed relatively high levels of academic engagement during baseline, resulting in a potential ceiling effect. In Bruhn et al. (2016), the use of SCORE IT was associated with a decrease in disruptive behavior and an increase in academic engagement for two participants. However, the authors noted that students were participating in a remedial reading intervention during periods in which the SCORE IT intervention was not in place, which may have contributed to improvements in academic engagement. Bruhn et al. (2017) evaluated effectiveness of SCORE-IT on the disruptive and academically engaged behaviors for three participants and observed variable levels of disruptive and academically engaged behavior across phases for all three participants. Similar to Bruhn et al. (2016), participants showed relatively low levels of disruptive behavior during baseline, making the demonstration of an intervention effect difficult.

Potential Advantages of Classroom Behavior Management (CBM) Apps

Besides potentially facilitating positive behavior change for students, CBM apps may have other advantages for teachers. CBM apps may be an efficient, automated tool for data collection and monitoring progress (Krach et al., 2017; Robacker et al., 2016; Wolf, 2015) as evidenced by results from preliminary research. CBM apps may eliminate the inaccuracies and unreliability of hand-calculated data (King et al., 2013; LeBel et al., 2013), which is time-consuming and can be prone to recording inaccuracies (Bellack & Hersen, 1998). Such technologies may allow data to be collected with high levels of integrity (Elswick et al., 2016), and could help teachers individualize interventions to suit student needs (e.g., Bruhn et al., 2020). In addition, technology-based interventions are generally viewed favorably by end-users (e.g., Bedesem, 2012; Szwed & Bouck, 2013), with high teacher acceptability reported in multiple studies (e.g., Christ & Christ, 2006; Kraemer et al., 2012; Radley et al., 2016). This perceived value and social acceptability may support the use of CBM apps in data-based problem-solving and intervention in schools.

Purpose

Despite the popularity and potential benefits of CBM apps, their effectiveness is relatively unknown (Dhir et al., 2013). It is interesting that researchers have identified teachers' perceptions (e.g., Cetin & Cetin, 2018) of the effectiveness of technologies as predictors of uptake, and teacher perceptions played a key role in their decision to introduce new technologies into classrooms (Blackwell et al., 2013; Sugar et al., 2004). However, perceptions of effectiveness must be accompanied by empirical demonstrations of positive behavior change for students associated with their use. Therefore, an examination of CBM apps is needed to ascertain their impact on effecting positive behavior changes in education settings. Because this is an emerging field and CBM apps' impact on encouraging positive behaviors is not well understood, the following systematic literature review aimed to answer the following research questions:

(1). Were CBM apps effective in supporting positive behavior change for school-aged children?
(2). What were the intervention components that may have supported positive behavior changes?
(3). What was the quality of included studies and the strength of the evidence for CBM apps?

Method

This review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A systematic search of the literature to identify articles for inclusion in this review was conducted in January 2020. The search focused on articles published between the dates of January 2007 to January 2020, with the 2007 selected as the start point because it marked the advent of the first mobile app with the release of the Apple iPhone (Apple Inc., 2007). A keyword search was conducted in seven databases: PsycINFO, Scopus, A+ Education, EBSCOhost, Taylor & Francis, ProQuest and ERIC. Table 1 provides an overview of the key words used to conduct the search. We used search characters "?" and "*" to account for alternate spellings of key words.

Inclusion Criteria

To be included, articles must have evaluated an intervention conducted in a school setting that included a CBM app. A CBM app was defined as a digital software application designed for mobile devices (phone or tablet) that provided teachers with specific features to support the assessment, monitoring, and delivery of interventions related to student behaviors in a classroom setting. Specific features included those that allowed for goal setting, prompting, self-monitoring, feedback (corrective or positive), data collection, and data analysis. Non-face-to-face environments (e.g., online learning environments) were excluded. The target population included school-aged participants between 5 and 18 years of age, inclusive of elementary, middle, and high schools but excluding university or preschool students. Studies evaluating non-app-based technologies (e.g., clickers, video modeling) were excluded. To be included, studies needed to provide a direct and objective measure of student behavior change. Behavior measured for change could include disruptive behavior (e.g., talking out

Table 1 Key Search Terms

Concepts	Key search terms
Classroom behavior management	("behavio?r management" OR "classroom management" OR "classroom behavio?r management" OR "behavio?r problems" OR "school based intervention" OR "behavio?r modification" OR "psychology education" OR "classroom behavio?r modification" OR "classroom behavio?r" OR "behavio?r analysis") AND
Student outcomes	("learning behavio?r*" OR "student outcome*" OR "student behavio?r*" OR "learning outcome*" OR "behavio?r" OR "academic achievement" OR "learning" OR "student attitudes" OR "classroom behavio?r" OR "behavio?r problems" OR "teacher student interaction*") AND
Apps	("technology" OR "digital technology" OR "mobile technology")

? and * denote search characters used to accommodate for alternate spelling of certain words during search for experimental studies

of turn, leaving seat) and academic engagement (e.g., maintaining attention, completing work) or achievement (e.g., learning outcomes in specific subjects). Studies that only measured teacher or parent outcomes (e.g., parent or teacher engagement, teacher behavior) were excluded. Studies that only used qualitative research methods, such as interviews and focus groups, to assess participant's perceptions of CBM apps were excluded. Theses and dissertations were excluded as we were specifically seeking to include articles that had been peer reviewed.

Screening Procedure

The initial screening identified 263 articles, which were subsequently uploaded to Covidence for review. Duplicates were removed, leaving 256 for further review (Fig. 1). The first author (a PhD student) screened the titles and abstracts of all 256 articles against the inclusion criteria to assess suitability for further review. This process resulted in the exclusion of 235 articles. The remaining 21 articles were read by the first author in their entirety. Of these, 15 articles met the inclusion criteria and were retained for full text review. A manual search of the reference list of each included article was then conducted by the first author to identify additional research that might have been missed in the initial search, but no additional articles were identified. Additional manual searches (e.g., page-by-page examination of the entire contents of specific journal issues) were not conducted because the ancestral search of included articles did not yield any additional articles for inclusion.

Interrater Reliability

The fifth author (a PhD student) independently screened the titles and abstracts of all 256 articles against the inclusion criteria to assess suitability for further review. Interrater reliability (IRR) was calculated for the title and abstract screening by dividing the total number of agreements between raters with the total number of agreements plus disagreements multiplied by 100. Total IRR for the title and abstract screening was 97%. An independent reviewer (second, third, or fourth author) independently reviewed the full text of the remaining 21 articles. Total IRR for the full text review was 82%. Disagreements between reviewers were then resolved through a consensus meeting between the first author and independent reviewer. Following the consensus meeting, all disagreements were resolved, and 100% agreement was established.

Data Extraction

A coding sheet was developed to permit the extraction of data from each included study on the following characteristics: (1) study information (authors, year, and country); (2) setting; (3) student participant characteristics; (4) study design; (5) intervention (independent variable); and (6) dependent variables. The first author coded all included articles.

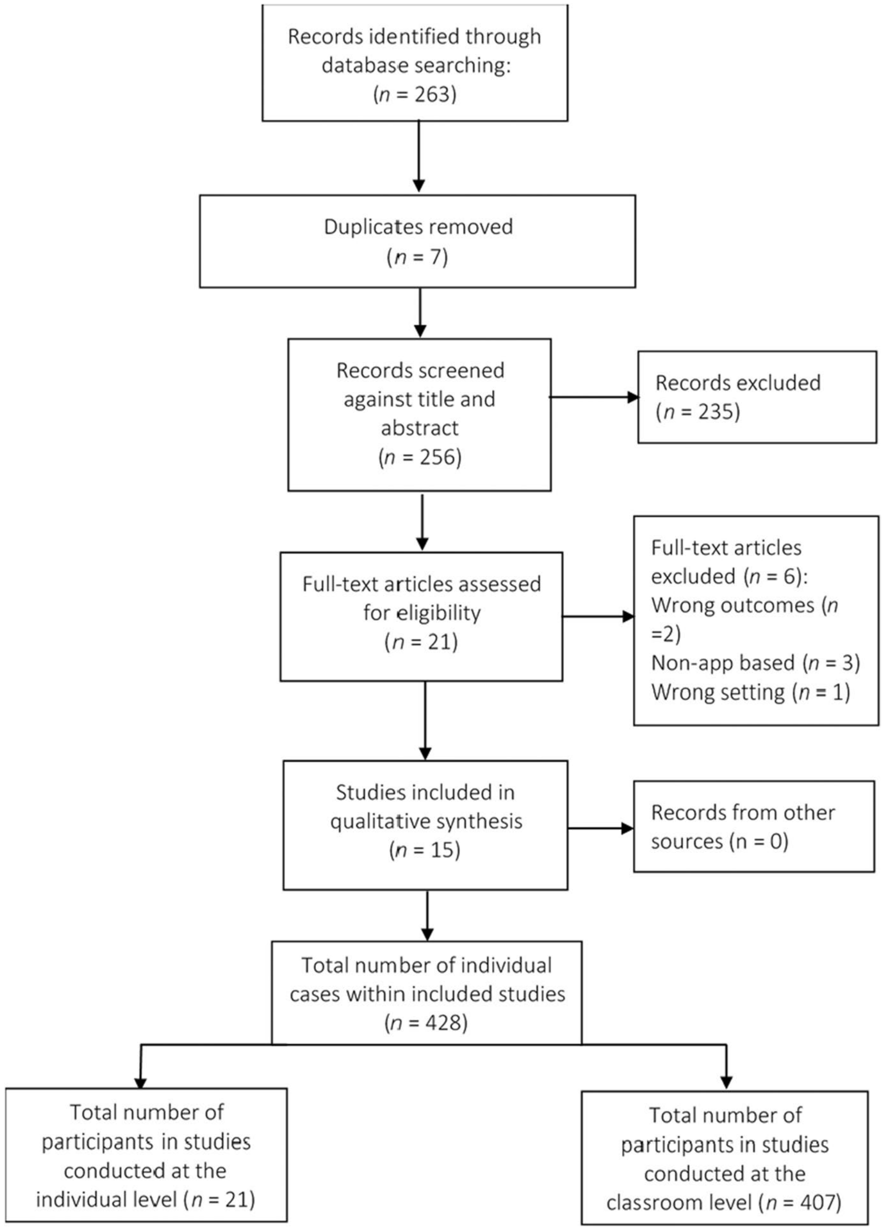

Fig. 1 Diagram Illustrating Identification and Screening Stages of Included Articles

Evaluation of Study Quality

Initial Quality Appraisal

The quality of the included studies (N = 15) was assessed using the What Works Clearinghouse (WWC) design standards version 5.0 (WWC, 2022). The first author first read the WWC Handbook and review protocols. Then, the first author conducted the initial quality evaluation for all included studies. First, all 15 studies were screened to determine if each case within each study (1) *Meet Standards without Reservations*, (2) *Meet Standards with Reservations*, or (3) *Does Not Meet Standards*. Studies that used single case designs (SCDs; n = 12) were reviewed to determine if: (1) data were presented in graphical or tabular format; (2) the independent variable(s) was systematically manipulated; (3) the dependent variable(s) was repeatedly measured, with IOA collected for a minimum number of data points in each phase and having met minimum IOA thresholds; (4) residual treatment effects were controlled for; and (5) there were a sufficient number of data points per phase. Group design studies ($n = 3$) were reviewed to determine if (1) groups were randomly assigned; (2) groups had low attrition (i.e., minimal dropouts of participants from a study); and (3) confounding factors were controlled for. If all criteria were satisfied, the study was eligible for a designation of (1) *Meet Standards without Reservations* or (2) *Meet Standards with Reservations*. If any of these criteria were not satisfied, the study was categorized as *Does Not Meet Standards*. No group design studies included in the current review satisfied the review criteria.

Risk of Bias

To assess risk of bias, Nonoverlap of All Pairs (NAP) was calculated for baseline trend and reversibility for each individual case in studies that were initially categorized as *Meet Standards without Reservations* or *Meet Standards with Reservations*. A “case” was defined as the unique pairing of an individual (or group) with a specific behavior of interest. Thus, a single participant for whom two different behaviors were measured constituted two cases. This operational definition allowed for a nuanced exploration of each behavior within and across individuals or groups, emphasizing the behavior’s role as a central unit of analysis. NAP is a statistical technique used to assess the degree of separation or nonoverlapping of data points in paired sets (e.g., baseline and intervention), and is the preferred method of calculating effect sizes for SCDs undergoing quality evaluation using the WWC design standards version 5.0. NAP was calculated by hand for each dependent variable for each case by the first author. A second independent rater (a PhD level behavior analyst and doctoral student supervisor familiar with WWC design standards) calculated NAP for 42% of included cases ($n = 11$). The second rater used the single-case effect size calculator (Pustejovsky et al., 2023) with the open-source R statistical software (version 4.3.1; R Core Team, 2023) to generate NAP values, and compared these values to the hand-calculated NAP scores obtained by the first author.

First, the initial baseline phase was reviewed to ensure that there was minimal therapeutic trend. To calculate NAP for baseline trend, all data points in the initial baseline phase except for the last three data points were compared to each of the last three data points for the initial baseline phase. When a data point in the last three observations showed improvement (evidence of a therapeutic trend) when compared to each initial baseline data point, it was given a score of 1. When a data point in the last three data points was equal to an initial baseline data point, it was given a score of 0.5. When a data point in the last three data points showed no improvement (evidence of a countertherapeutic trend) when compared to each initial baseline data point, it was given a score of 0. The sum of all pairwise comparisons was calculated and then divided by the total number of comparisons to yield a NAP score. Cases with scores less than 0.85 showed limited risk of bias and were eligible to receive a rating of *Meet Standards without Reservations.* Cases with scores above 0.85 showed potential risk of bias (evidence of improvement in behavior during baseline) and were eligible to receive a rating of *Meet Standards with Reservations*. IRR was 81.8% for baseline trend.

Next, the data in the second baseline phase were compared to data in the initial baseline phase to assess reversibility. Each initial baseline data point was individually compared to each data point in the reversal to baseline phase. The calculations and scoring were identical to that described above. Cases with scores less than 0.85 showed limited risk of bias and were eligible to receive a rating of *Meet Standards without Reservations.* Cases with scores above 0.85 showed potential risk of bias (limited evidence of reversibility) and were eligible to receive a rating of *Meet Standards with Reservations*. IRR was 90.9% for reversibility.

Effect Size Indicators

Finally, NAP was calculated to determine the proportion of data in the intervention phase that demonstrated improvement over the baseline phase. Each baseline data point was individually compared to each data point in the subsequent intervention phase. The calculations and scoring were identical to that described above. High NAP values, typically above 0.92, indicate strong evidence of an intervention effect. NAP values in the middle range (0.66 and 0.92) are indicative of moderate evidence of an effect.

Table 2 Description of All Studies which Met Inclusion Criteria for the Current Systematic Review

References	Country	School setting	Participant characteristics	Study design	Intervention/Type	Dependent variables
Beckman et al. (2019)	USA	Upper elementary, special education	*N* = 2 M, Grades 5 & 6	ABAB with withdrawal	I-Connect (Self-monitoring)	On-task behavior Academic outcomes
Bruhn et al. (2015)	USA	Middle, inclusive classroom for struggling readers	*N* = 2 M = 1, F = 1 Grades 7 & 8	ABBAB (Don) ABAB (Jess)	SCORE IT (Self-monitoring)	AE (for both) DB (Don) TO (Jess)
Bruhn et al. (2016)	USA	Middle, noncategorical special education	*N* = 2 M, Grades 6 & 7	ABAB + maintenance	SCORE IT (Self-monitoring)	AE DB
Bruhn et al. (2017)	USA	Public middle	*N* = 3 M = 1, F = 2 Grades 6 & 7	MBL across participants	SCORE IT (Self-monitoring)	AE DB
Bunch-Crump & Lo (2017)	USA	Public elementary	*N* = 1 M Grade 3	MBL with reversal	CICO vs I-Connect (FBSM) (Self-monitoring)	AE DB
Chen et al. (2017)	China	Primary	*N* = 124 Grade 1	Pre-post	CBMS (Token Economy)	Academic Achievement
Chiarelli et al. (2015)	USA	Primary	*N* = 24, Grade 1 M = 16, F = 8	Pre-post	ClassDojo (Token Economy)	Positive & negative behaviors
Clemons et al. (2016)	USA	Public high, general education	*N* = 3 M = 2 (Grade 11) F = 1 (Grade 9)	ABAB withdrawal	I-Connect (Self-monitoring)	On-task behavior
Crutchfield et al. (2015)	USA	Middle, special education	*N* = 2 M, both 14yrs	ABAB with MBL	I-Connect (Self-monitoring)	Stereotypic behavior
Dillon et al. (2019)	USA	Middle (*n* = 2)	*N* = 74 M = 35, F = 39 Grades 5	ABAB withdrawal	Tootling with ClassDojo (Token Economy)	AE DB
Homer et al. (2018)	Hong Kong	Elementary	*N* = 120, M = 71, F = 49 Grades 1 to 4	Pre-post	ClassDojo vs. nondigital token economy (Token Economy)	Student learning Student behavior (positive & negative)
Lynne et al. (2017)	USA	K–8 Primary, general education	*N* = 65 M = 33, F = 32 Grades 1 and 4	ABAB withdrawal	GBG with ClassDojo (Token Economy)	DB AE
Rosenbloom et al. (2016)	USA	Elementary, general education	*N* = 1 M, Grade 3	ABAB withdrawal	I-Connect (Self-monitoring)	On-task behavior DB
Vogelgesang et al. (2016)	USA	Elementary, general education	*N* = 3 M = 1, F = 2 Grade 5	ABAB + maintenance	SCORE IT (Self-monitoring)	AE
Wills & Mason (2014)	USA	High, general education	*N* = 2 M, Grade 9	ABAB withdrawal	I-Connect (Self-monitoring)	On-task behaviors DB (gen)

AE academic engagement, *CBMS* classroom behavior management system, *CICO* check in check out, *DB* disruptive behavior, *F* female, *FBSM* function based self-monitoring, *GBG* Good Behavior Game, *gen* generalization, *m/M* male, *TO* talk outs, *yrs* year

Low NAP values (< 0.66) suggest weak or no evidence of an effect (Parker & Vannest, 2009). IRR was 100% for intervention effect sizes.

Results

Table 2 provides an overview of the characteristics of the included studies.

Location and Settings Characteristics

Most studies were conducted in the United States (*n* = 13), with one study conducted in China (Chen et al., 2017) and another in Hong Kong (Homer et al., 2018). All studies were conducted within the student participants' schools and respective classrooms. Interventions were delivered in special education classrooms (*n* = 4) and general education classrooms (*n* = 5), whereas information about classroom type was missing from six studies (Bruhn et al., 2017; Bunch-Crump & Lo, 2017; Chen et al., 2017; Chiarelli et al., 2015; Dillon et al., 2019; Homer et al., 2018).

Participant Characteristics

In the included studies, participants (*n* = 428) ranged in age from 6 to 17 years. Most participants attended elementary schools and were in grades 1 to 6 (*n* = 414). Few studies included middle school students from grades 6 to 8 (*n* = 9) or high school students (*n* = 5), and those that did were in grades 9 and 11 only. Some studies did not provide information on certain student demographics. Ages were not reported in six studies (Bruhn et al., 2016; Bruhn et al., 2017; Chen et al., 2017; Chiarelli et al., 2015; Dillon et al., 2019; Lynne et al., 2017) with researchers reporting grade level instead. Participants represented ethnically diverse groups, including Chinese (*n* = 244), white (*n* = 71), African American (*n* = 33), Hispanic (*n* = 8), Middle Eastern (*n* = 4), multiethnic (*n* = 3), and Native American (*n* = 1). Ethnicity of specific students was not reported in two studies (Bruhn et al., 2017; Lynne et al., 2017). There were more male (*n* = 170) compared with female (*n* = 134) participants. In the majority of studies (*n* = 12), participants had one or more disabilities that included, but not limited to, specific learning disabilities, attention deficit hyperactivity disorder, autism, and intellectual disability. One study (Bunch-Crump & Lo, 2017) included a neurotypical student with problem behaviors. Diagnostic information was not provided in two studies (Chen et al., 2017; Homer et al., 2018).

Dependent Variables and Outcome Measures

Dependent variables measured in the studies included disruptive behavior (*n* = 8), academic engagement (*n* = 7), on-task behavior (*n* = 4), positive (*n* = 2) and negative behaviors (*n* = 2), talk-outs (*n* = 1), and stereotypic behaviors (*n* = 1). Only a few studies (*n* = 3) measured specific learning outcomes following the intervention. For example, Beckman et al. (2019) measured academic outcomes in math and written expression, Chen et al. (2017) measured academic achievement across three semesters in a Chinese language class, and Homer et al. (2018) measured student learning in English reading and speaking classes. Researchers used various methods to measure changes in the dependent variables, including whole interval recording (*n* = 5), partial interval recording (*n* = 5), momentary time sampling (*n* = 3), frequency (*n* = 3), duration recording (*n* = 2), summative assessment (*n* = 3), and a behavior chart (*n* = 1) to record the occurrences of behaviors. Only two studies analyzed data that was automatically generated by the CBM app (Chen et al., 2017; Chiarelli et al., 2015).

Research Designs

Twelve studies used SCD methodology to demonstrate effect of the intervention (independent variable) on the target behavior (dependent variable) (Beckman et al., 2019; Bruhn et al., 2016; Bruhn et al., 2015; Bruhn et al., 2017; Bunch-Crump & Lo, 2017; Clemons et al., 2016; Crutchfield et al., 2015; Dillon et al., 2019; Lynne et al., 2017; Rosenbloom et al., 2016; Vogelgesang et al., 2016; Wills & Mason, 2014). The majority of these studies (*n* = 10) used the reversal/withdrawal design or variation thereof (e.g., ABAB with maintenance, ABAB with embedded multiple baseline (MBL) across participants). Two studies (Bruhn et al., 2017; Bunch-Crump & Lo, 2017) used an MBL across participants design. The remaining three studies used group designs and analyzed data using regression (Chen et al., 2017), independent *t*-tests (Homer et al., 2018), or pre- and post-outcome measures (Chiarelli et al., 2015).

Types of Classroom Behavior Management (CBM) Apps

The CBM apps identified in this review fit into two broad intervention categories: self-monitoring and token economy. Ten studies evaluated the self-monitoring apps I-Connect ($n = 6$) and SCORE IT ($n = 4$), four studies examined the effectiveness of ClassDojo, and one evaluated the Classroom Behavior Management System (CBMS) app. The I-Connect and SCORE IT apps are similar in that they allow students to self-monitor their behavior. Using the SCORE IT app, students are able to rate their behavior at specific intervals using a Likert-type rating scale. Teachers are also able to rate students' behaviors using the same scale, and then compare their ratings with those of students. In addition, the SCORE IT app allows teachers and students to set individual behavior goals and track student progress toward goals using graphs generated by the app. Using the I-Connect app, students are prompted to self-monitor and self-record their own behavior at specific intervals by responding to a "yes/no" question (e.g., "Are you on task?"). Similar to SCORE IT, I-Connect allows students and teachers to create individual behavior goals and provides graphic displays of student data. Both SCORE IT and I-Connect are primarily designed to be used by individual students.

ClassDojo is an educational technology platform that helps teachers manage and improve classroom behavior and communication with students and parents. It offers a variety of features to support teachers in rewarding positive student behavior. Teachers can award digital points to individual students or groups of students for demonstrating positive behaviors, such as participation, teamwork, completing assignments, or being respectful. The point system can be customized by teachers to align with the teacher's classroom behavior expectations. ClassDojo can also provide a visual representation of each student's progress by displaying awarded points on a screen or smartboard in the classroom. Teachers can also share data with and message parents using the ClassDojo app. Chen et al. (2017) developed a custom app, the CBMS, for use in their study. The design of CBMS was based on ClassDojo and allowed teachers to create individual avatars for each student and rate student behavior by clicking the avatar. Positive ratings can be delivered for behaviors defined by the teacher, such as on-task behavior or correct responding. Negative ratings can be delivered for undesirable behaviors. Similar to ClassDojo, CBMS can provided a visual representation of rating data for individual students and be used by teachers to facilitate data-based decision making or shared directly with students and parents. In addition, summarized results can also be shown to the entire class or an individual student through a tablet during the class.

Antecedent-Based Strategies

The six studies that evaluated I-Connect included antecedent intervention strategies such as behavior goal setting and prompting. All six studies incorporated goal setting, with teachers deciding on behavior goals before implementing the CBM intervention. In several studies, the I-Connect app prompted students to self-monitor by responding to a question such as "Am I on task?" (Beckman et al., 2019; Clemons et al., 2016; Rosenbloom et al., 2016; Wills & Mason, 2014), "Quiet hands and mouth?" (Crutchfield et al., 2015) or "Following the rules?/Need help?" (Bunch-Crump & Lo, 2017) at specific time intervals by touching either "Yes" or "No" on the screen. In three of the six I-Connect studies (Crutchfield et al., 2015; Rosenbloom et al., 2016; Wills & Mason, 2014) the app was evaluated in the absence of programmed consequences (such as positive reinforcement or response cost) for student on-task or disruptive behavior. Rosenbloom et al. (2016) and Wills and Mason (2014) showed that the use of the app was associated with an increase in individual students' on-task behaviors and a decrease in disruptive behaviors. Likewise, Crutchfield et al. (2015) showed that individual students' academic engagement increased and stereotypic behaviors decreased following the introduction of I-Connect. However, in two of the three studies (Crutchfield et al., 2015; Wills & Mason, 2014) students were also prompted by the teacher or researcher to respond to the question if they did not respond to the prompt provided by the I-Connect app within 3 s. Therefore, the degree to which the use of the app alone, in the absence of teacher-mediated prompts, was effective for increasing on-task behavior and decreasing disruptive behavior is unknown.

In all four studies that evaluated the SCORE IT app (e.g., Bruhn et al., 2016; Bruhn et al., 2015; Bruhn et al., 2017; Vogelgesang et al., 2016), antecedent-based intervention strategies such as prompting and goal setting were used. Vogelgesang et al. (2016) examined the effectiveness of the SCORE IT app in the absence of programmed consequences for accurate self-monitoring or appropriate behavior and found that the introduction of SCORE IT was associated with an increase in academic engagement, with gains maintained at 2- and 4-week follow-ups.

Antecedent-Based Strategies Combined with Consequence-Based Strategies

In two of the six I-Connect studies (e.g., Beckman et al., 2019; Bunch-Crump & Lo, 2017) researchers introduced a minimum target behavior goal for reinforcement. For example, positive reinforcement (e.g., praise, rewards) was provided if the participant met at least 80% of their behavior goals. In three out of the six studies (e.g., Beckman et al., 2019; Bunch-Crump & Lo, 2017; Clemons et al., 2016), the effectiveness of the I-Connect app paired with consequence-based intervention strategies such as positive reinforcement and corrective feedback was evaluated. For instance, participants in Bunch-Crump and Lo (2017) received socially mediated reinforcement in the form of praise, acknowledgement to questions answered, pats on the back or a thumbs up for academic engagement. In addition to teacher praise, participants in two studies (Beckman et al., 2019; Clemons et al., 2016) received behavior specific feedback and corrections when engaging in an off-task behavior, and received back-up reinforcers (i.e., access to preferred items) for meeting their behavior goal (e.g., 80% of intervals of on-task behavior). However, results from the studies that paired the use of I-Connect with consequence interventions were varied. Beckman et al. (2019) showed that on task behaviors increased for both participants, but academic outcomes (e.g., math test scores) improved for only one out of two students. Clemons et al. (2016) showed that on task behaviors increased and maintained at follow-up following the introduction of I-Connect paired with consequence-based intervention components.

The effectiveness of the SCORE IT app when paired with consequence strategies was examined in three (Bruhn et al., 2016; Bruhn et al., 2015; Bruhn et al., 2017) of the four studies. In all three studies, participants received teacher praise and back-up reinforcers (i.e., access to preferred items) for engaging in appropriate behavior and meeting their behavior goals and received behavior specific feedback and corrections for engaging in an off-task behavior. In Bruhn et al. (2016), the use of the SCORE IT app was associated with a decrease in disruptive behavior and an increase in academic engagement for two participants. Visual analysis of the graphed data indicated that both students' behavior changed when SCORE-IT paired with both antecedent strategies (prompting to self-monitor) and consequence strategies (praise for appropriate behavior and corrections for off-task behavior) was introduced. However, in Bruhn et al. (2015) there was a decrease in disruptive behavior for both participants but an increase in academic engagement for only the first participant. It should be noted that it is unclear if one participant was unable to meet her goal for academic engagement because of high variability and overlapping data points across phases. In Bruhn et al. (2017), the effectiveness of SCORE-IT on the disruptive and academically engaged behaviors for each participant were unclear due to variability in the data across phases and methodological design flaws (e.g., less than three attempts to demonstrate an intervention effect using a multiple baseline design).

Antecedent-Based Intervention Strategies Combined with Check In/Check Out (CICO)

In a unique example, Bunch-Crump and Lo (2017) evaluated the effects of Check-In Check-Out (CICO) with four participants. However, the data showed that CICO alone was not effective for reducing one participant's disruptive behavior. As a result, the intervention was individualized and intensified further for this participant using I-Connect. During the CICO plus I-Connect phase, the app prompted the participant to self-monitor whether he was following the classroom rules at specific time intervals and provided a reminder to the participant to request help from the teacher every 10 min. When I-Connect was introduced, there was a small decrease in disruptive behavior. However, few data points and overlapping data between the CICO alone and the CICO plus I-Connect phases made it difficult to determine the effectiveness of the addition of I-Connect to CICO.

Consequence-Based Strategies

Two correlational studies (Chiarelli et al., 2015; Homer et al., 2018) showed positive results associated with the use of the Class Dojo app. In both studies, the use of ClassDojo involved setting a predetermined goal for all students in the classroom and reinforcement (a consequence-based intervention) for specific desired behaviors. Participants were awarded Dojo points for positive behaviors (Chiarelli et al., 2015), and for their learning and behavior goals (Homer et al., 2018). In addition, a response cost system was also implemented in both studies whereby points were deducted when students engaged in negative behaviors. Chiarelli et al. (2015) evaluated a CBM intervention delivered via the ClassDojo app. Students chose a personalized avatar prior to commencing the study to represent them on the app, and each student received individualized feedback based on their behavior throughout the study. Students received auditory feedback in the form of a ring sound when they engaged in positive behavior and a buzz sound when they engaged in undesirable classroom behaviors. In addition, a Dojo point (e.g., token) was awarded to each student who was engaging in positive behavior when the ring sounded and a Dojo point was deducted (response cost) for each student who was engaging in undesirable classroom behavior when the buzz sounded. The authors found that students engaged in more positive behaviors and less undesirable behaviors following the introduction of the ClassDojo intervention.

Homer et al. (2018) conducted a comparison of student behavior in classrooms implementing CBM with and without ClassDojo. In the control group, the teacher implemented a paper and pencil point system. Points were awarded to students who displayed positive behaviors (positive reinforcement) and deducted from students who displayed inappropriate behaviors (response cost). In the treatment group, the same point system was implemented except the ClassDojo app was used to facilitate the delivery and deduction of points. In both groups, students were able to exchange their earned points for a backup reinforcer (e.g., new badges/avatars and stationery). The authors found the Grade 3–4 treatment group students showed statistically significantly higher oral (language) post test scores compared to Grade 3–4 control group participants at the conclusion of the intervention. However, no significant difference in reading post-test scores between treatment and control groups were observed for grades 1–2 students. Based on collected data, Homer et al. (2018) found that overall students in the experimental group behaved better than students in the control group as a result of teachers using ClassDojo. Students who experienced both types of CBM interventions (with and without ClassDojo) were surveyed at the end of the study to gather information about student perceptions of the intervention. Students who experienced the ClassDojo intervention reported more positive perceptions of the intervention than students in the control group.

Two studies evaluated ClassDojo in conjunction with other intervention components. Dillon et al. (2019) used ClassDojo with positive peer reporting known as Tootling (Skinner et al., 2000), which encouraged students to monitor and report on the prosocial behaviors of their classmates (as opposed to "tattling"). In addition to peer reporting, students received verbal praise from their teachers and were permitted to exchange their points for back-up reinforcers (e.g., extra free time or small edible items) contingent on academic engagement. Dillon et al. showed that the introduction of Tootling with ClassDojo was associated with an increase in students' academic engagement and a decrease in disruptive behavior across three classrooms. Lynne et al. (2017) used ClassDojo paired with the Good Behavior Game (GBG; Barrish et al., 1969), an interdependent group contingency management program in which points are allocated to teams (rather than individual students) exhibiting appropriate behaviors. In this study, teams that earned a specific number of points for academic engagement were permitted to exchange their points for back-up reinforcers (e.g., tickets and or candy). Results showed that GBG paired with ClassDojo was associated with an increase in students' academic engagement and a decrease in disruptive behavior in two out of three classrooms.

Chen et al. (2017) demonstrated that the Classroom Behavior Management System (CBMS) app was effective in improving students' academic achievement by enhancing their positive classroom

Table 3 Nonoverlap of All Pairs for Each Case in the Single Case Design Studies Eligible for Meet WWC Standards without/with Reservations

References	Cases	Intervention	Dependent variable	Design	Baseline trend (NAP)	Reversibility (NAP)	Effect size estimate	WWC design standards
Beckman et al. (2019)	Cody	I-Connect	On-task behavior	ABAB	0.50	0.52	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Cody	I-Connect	Academic outcome (Math)	ABAB	0.25	0.60	A1B1 = 0.93, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Brian	I-Connect	On-task behavior	ABAB	0.38	0.58	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Brian	I-Connect	Academic outcome (writing)	ABAB	0.38	0.58	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
Bruhn et al. (2015)	Jess	SCORE IT	Academic engagement	ABAB	0.38	0.31	A1B1 = 0.60, weak effect A2B2 = 0.70, moderate effect	Meets WWC Standards Without Reservations
	Jess	SCORE IT	Talk-outs	ABAB	0.06	0.47	A1B1 = 0.89, moderate effect A2B2 = 0.80, moderate effect	Meets WWC Standards Without Reservations
Clemons et al. (2016)	Keith	I-Connect	On-task behavior	ABAB	0.28	0.77	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Brad	I-Connect	On-task behavior	ABAB	0.44	0.70	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Miranda	I-Connect	On-task behavior	ABAB	0	0.84	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
Dillon at al. (2019)	Classroom A ($n = 35$)	Tootling via ClassDojo	Academically engaged behavior	ABAB	0	0.09	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Classroom A ($n = 35$)	Tootling via ClassDojo	Disruptive behavior	ABAB	0.17	0	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations

Table 3 (continued)

References	Cases	Intervention	Dependent variable	Design	Baseline trend (NAP)	Reversibility (NAP)	Effect size estimate	WWC design standards
	Classroom B ($n = 20$)	Tootling via ClassDojo	Academically engaged behavior	ABAB	0.75	0.73	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Classroom B ($n = 20$)	Tootling via ClassDojo	Disruptive behavior	ABAB	0.63	0.90	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards With Reservations
	Classroom C ($n = 19$)	Tootling via ClassDojo	Academically engaged behavior	ABAB	0.34	0.73	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Classroom C ($n = 19$)	Tootling via ClassDojo	Disruptive behavior	ABAB	0.33	0.72	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
Lynne et al. (2017)	Classroom A ($n = 27$)	ClassDojo with GBG	Disruptive behavior	ABAB	0	0.94	A1B1 = 0.96, strong effect A2B2 = 0.98, strong effect	Meets WWC Standards With Reservations
	Classroom A ($n = 27$)	ClassDojo with GBG	Academically engaged behavior	ABAB	0	0.64	A1B1 = 0.87, moderate effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Classroom B ($n = 19$)	ClassDojo with GBG	Disruptive behavior	ABAB	0.03	0.58	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Classroom B ($n = 19$)	ClassDojo with GBG	Academically engaged behavior	ABAB	0.13	0.40	A1B1 = 1, strong effect A2B2 = 0.92, moderate effect	Meets WWC Standards Without Reservations
	Classroom C ($n = 19$)	ClassDojo with GBG	Disruptive behavior	ABAB	0.50	1	A1B1 = 1, strong effect A2B2 = 0.87, moderate effect	Meets WWC Standards With Reservations
	Classroom C ($n = 19$)	ClassDojo with GBG	Academically engaged behavior	ABAB	0.25	164	A1B1 = 0.97, strong effect A2B2 = 0.98, strong effect	Meets WWC Standards Without Reservations
Rosenbloom et al. (2016)	Student 1	I-Connect	On-task behavior	ABAB	0.67	0.64	A1B1 = 0.98, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations

Table 3 (continued)

References	Cases	Intervention	Dependent variable	Design	Baseline trend (NAP)	Reversibility (NAP)	Effect size estimate	WWC design standards
Vogelgesang et al. (2016)	John	SCORE IT	Academic engagement	ABAB	1	1	A1B1 = 1, strong effect A2B2 = 1, strong effect	Meets WWC Standards With Reservations
Wills & Mason (2014)	Student 2	I-Connect	On-task behavior	ABAB	0.03	0.67	A1B1 = 0.98, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations
	Student 2	I-Connect	Disruptive behavior	ABAB	0	0.53	A1B1 = 0.71, moderate effect A2B2 = 0.68, moderate effect	Meets WWC Standards Without Reservations
	Student 1	I-Connect	Disruptive behavior	ABAB	0.07	0.66	A1B1 = 0.98, strong effect A2B2 = 1, strong effect	Meets WWC Standards Without Reservations

ABAB = reversal/withdrawal design, *A1B1* = initial baseline phase followed by intervention phase, *A2B2* = return to baseline followed by second intervention phase, *GBG* = Good Behavior Game, *WWC* = What Works Clearinghouse

behaviors. Positive behavior ratings explained 38.46% of the variance in students' posttest scores ($p < .001$) across three semesters (Chen et al., 2017). In this study, students were awarded points for academic achievement and a response cost system was also implemented whereby points were deducted when they engaged in negative behaviors. However, Chen et al. did not document whether participants were able to exchange their points for back-up reinforcers.

Quality of Included Studies

The quality of the research conducted to date on CBM apps was assessed by applying the WWC design standards version 5.0 (WWC, 2022). Out of a total of 428 participants from single case design and group studies, 148 participants from eight studies (Beckman et al., 2019; Bruhn et al., 2015; Clemons et al., 2016; Dillon et al., 2019; Lynne et al., 2017; Rosenbloom et al., 2016; Vogelgesang et al., 2016; Wills & Mason, 2014) met the WWC design standards (see Table 3). These eight studies were then subsequently reviewed for the presence or absence of additional features (e.g., IOA, fidelity, generalization, maintenance, and social validity data). For these individual cases, effect size indicators were calculated using NAP.

Reliability, Fidelity, Social Validity, Maintenance and Generalization

Acceptable levels of interobserver agreement (IOA; range: 80%–100%), indicating a high degree of reliability was reported in the eight studies where 148 participants were either rated as *Meets WWC Standards without* or *with Reservations*. All but one of these eight studies (Beckman et al., 2019) reported measures of procedural fidelity. Procedural integrity scores were at or above 75% for the other nine studies.

Social validity assessment measures acceptability and satisfaction regarding the procedures of an intervention. Of the seven studies that measured social validity, researchers in four studies sought feedback from both teachers and students (Bruhn et al., 2015; Clemons et al., 2016; Rosenbloom et al., 2016; Wills & Mason, 2014), and researchers in three studies sought feedback from teachers only (Dillon et al., 2019; Lynne et al., 2017; Vogelgesang et al., 2016). Social validity was not measured in Beckman et al. (2019) because the first author was both the teacher and researcher which, it could be argued, would have limited the objectivity of the assessment. All seven studies reported positive perceptions. Teachers and students viewed the use of I-Connect, SCORE IT, Tootling with ClassDojo and GBG with ClassDojo favorably. Dillon et al. (2019) reported moderate to high social validity, although teachers noted some concerns about the durability of the effects of the intervention.

Maintenance data were only reported in two studies (Clemons et al., 2016; Vogelgesang et al., 2016). In Clemons et al. (2016), on-task behaviors were maintained at 2- and 4-week follow-ups while students were allowed to continue using I-Connect. In Vogelgesang et al. (2016), academic engagement was maintained at 2-, 3- and 4-week follow-ups when use of the SCORE IT app was faded. Generalization data was only reported in one study. Wills and Mason (2014) included disruptive behavior as a measure of generalization, to see if the introduction of I-Connect would result in a concomitant decrease in disruptive behavior as on-task behaviors increased. In their study, disruptive behavior decreased following the introduction of I-Connect, providing evidence of generalization.

Risk of Bias

Table 3 depicts the NAP scores for the 26 cases included in the eight studies categorized as *Meets Standards without Reservations* and *Meets Standards with Reservations*. Twenty-one cases from seven studies (Beckman et al., 2019; Bruhn et al., 2015; Clemons et al., 2016; Dillon et al., 2019; Lynne et al., 2017; Rosenbloom et al., 2016; Wills & Mason, 2014) showed limited risk of bias and were rated as *Meets WWC Standards without Reservations* (see Table 3). Five cases from three studies (Dillon et al., 2019; Lynne et al., 2017; Vogelgesang et al., 2016) showed some risk of bias and were rated as *Meets WWC Standards with Reservations*.

Effect Size Indicators

Table 3 also depicts effects size estimates for the 26 cases included in the eight studies categorized as *Meets Standards without Reservations* and *Meets Standards with Reservations.* All four studies that used I-Connect (Beckman et al., 2019; Clemons et al., 2016; Rosenbloom et al., 2016; Wills & Mason, 2014) showed evidence of moderate to strong intervention effects, with effect size estimate ranging from 0.68 (moderate) to 1 (strong). Effect size estimate (NAP) ranged from weak to strong (0.60 to 1) for studies that used SCORE IT. Strong effect size estimates were obtained for the study that evaluated Tootling plus ClassDojo (Dillon et al., 2019). For studies evaluating the GBG plus ClassDojo, effect size estimates ranged from moderate to strong (Lynne et al. (2017).

Discussion

Although mobile technologies such as apps have increased in popularity in many educational settings due to their ease of use and low cost (Kinash et al., 2012), evidence regarding their effectiveness for supporting positive behavior change in students has remained relatively unknown. Therefore, the aims of the present study were to systematically review the published research on CBM apps to evaluate whether they were effective for supporting positive behavior changes in school-aged children, to identify the intervention components that supported positive behavior change, and to assess the quality and strength of the evidence for the included studies using the WWC (2022) design standards. In the current study, we found that CBM apps were often used as part of antecedent interventions that involved teaching students to self-monitor their behavior or as part of consequence interventions that involved the delivery of positive reinforcement for desirable behavior. When used as part of self-monitoring interventions, apps such as I-Connect and SCORE IT prompted students to self-monitor their behavior at specific points in time using a rating scale within the app itself. When used as part of a class-wide reinforcement system, apps such as ClassDojo provided opportunities for students to create individual avatars within the app for teachers to award points or ratings to individual students for desirable behaviors.

Assessment of the quality of the included studies yielded mixed results. Only 8 out of 15 included studies (53%) met criteria to receive a WWC rating of *Meets Standards without Reservations* or *Meets Standards with Reservations,* with none of the group design studies meeting WWC standards. Of the eight studies that met standards with or without reservations, the majority of cases (69%) showed limited risk of bias. Moderate to strong effect sizes were observed for the majority of cases included in these eight SCD studies, suggesting that the use of CBM apps might promote positive behavior change for students when used with individual students.

Although the review indicated that CBM apps could promote positive student behaviors, apps were often used in combination with other behavior change strategies. For instance, evidence of effectiveness for ClassDojo was only evident when implemented alongside other interventions such as Tootling (Dillon et al., 2019) and the GBG (Lynne et al., 2017). These findings raise questions about the effectiveness of CBM apps as standalone intervention components, and about the added benefits of CBM apps when compared to similar behavior change strategies (such as self-monitoring and positive reinforcement) when delivered without the use of apps. It is possible that CBM apps could enhance the efficacy of other behavior change strategies by making the delivery of these interventions easier, more probable, or more manageable for teachers. For example, ClassDojo might make implementation of class-wide reinforcement systems easier by allowing the teacher to deliver reinforcers (Dojo points) to students more often or from afar, simply by selecting the avatar and awarding a point to the student in the app.

At present, there is limited evidence to demonstrate that the use of CBM apps alone, in the absence of other behavior change strategies, are effective for facilitating positive student behavior change. However, the findings of studies included in this review suggest that apps may be effective when paired with antecedent and/or consequence-based behavior change strategies, such as self-monitoring and positive reinforcement. In what follows, we provide preliminary recommendations for educators and other professionals who work in school contexts who may be considering the use of CBM apps.

Recommendations for Teachers and Other Professionals Working in Schools

First, we recommend that teachers, school psychologists, and behavior analysts consider using CBM apps as part of a comprehensive and well-designed CBM intervention package. To do so, teachers in particular may benefit from learning about the behavioral principles underpinning CBM apps, because the findings of this review show that they may be effective when combined with empirically validated behavior change strategies (Hammonds et al., 2013; Riden et al., 2019; Robacker et al., 2016). None of the studies included in this review described any training being provided to users in behavior principles or behavior change strategies prior to using the CBM app with students. By understanding the behavior principles that underpin CBM apps, teachers might be better informed to use the apps in ways that will be optimally effective.

Second, the current findings suggest that effectiveness of CBM apps can vary greatly depending on the specific app, the context in which it is used, and the student cohort. Thus, it is important for teachers and other professionals working within schools to carefully consider what type of app may be fit for purpose, depending on the strengths and needs of students and the context in which the app will be used. For example, ClassDojo may be an appropriate app for use in situations where the teacher wishes to identify and positively reinforce expected behaviors displayed by all students in the classroom. By contrast, SCORE IT may be more suitable for individual students who would benefit from more personalized support to demonstrate and self-monitor their use of appropriate behavior.

A key purported advantage of CBM apps is that they offer real time data collection and data analysis, which can provide important information to teachers, behavior analysts and school psychologists about student behavior and learning. In addition, the automation of some elements of CBM interventions, such as signaling to students when to self-monitor their behavior or providing behavior specific feedback to students, may help teachers to simultaneously deliver instruction and implement classroom behavior management programs (Beaman et al., 2007; Clunies-Ross et al., 2008; Ratcliff et al., 2010). This may reduce workload for teachers and make implementation of CBM strategies easier. In addition, the automation of data may provide real time information to teachers to allow them to adjust their teaching and/or use of CBM strategies based on student performance. Thus, we recommend that educators and professionals working in school settings consider using apps to facilitate the collection and analysis of data to inform data-based decision making. For example, the use of CBM apps may prove useful in identifying issues before they escalate which may facilitate the provision of early intervention supports for students who may be at risk. Automation of data collection may also facilitate data analysis and data-based problem solving, which have been identified as critical features of multitiered systems of support in school contexts (McIntosh & Goodman, 2016). However, teachers, behavior analysts and school psychologists should not rely exclusively on the information generated by CBM apps to make decisions about student learning and the need for other modifications. Information generated by CBM apps should supplement, rather than replace, other sources of data commonly collected and reviewed by school team members such as teacher observations and standardized assessment results.

Limitations and Future Research Directions

Some limitations of the current review should be acknowledged. First, CBM apps were implemented in combination with other behavior change strategies (such as self-monitoring, feedback, and/or back up reinforcement) in the majority of studies included in this review. Thus, it was difficult to draw conclusions about which components of CBM interventions, such as the use of the app, contributed to positive behavior change. Component analyses of CBM interventions that integrate the use of apps may allow researchers to identify the specific aspects of the treatment package that are most critical to the success of the intervention (Baer et al., 1968). Findings of component analyses could also inform app developers, researchers, behavior analysts, school psychologists and educators on how to enhance the effectiveness of CBM apps by clarifying ways to optimize their use, such as how to combine CBM apps with other evidence-based behavior change strategies.

The majority of studies that met inclusion criteria in the current review used SCD, and as such represent findings for a relatively small pool of participants. Mixing SCDs and group design studies made it difficult to synthesize results or draw unified conclusions. As research in this field is expanding, future reviews could adopt a meta-analytic methodology, which may offer a more nuanced understanding of effect sizes and overall trends. Moreover, expanding the search criteria to incorporate a greater variety of sources, such a grey literature or unpublished studies (e.g., theses, dissertations) might counteract any publication bias and present a more complete picture of the available evidence.

Third, few studies included data on maintenance and generalization of intervention effects. Generalization was only assessed in one study (e.g., Wills & Mason, 2014) and maintenance measured in only two studies (Clemons et al., 2016; Vogelgesang et al., 2016). Future research conducted in this field should aim to investigate generalization and maintenance of student behaviors at follow-up sessions because when a behavior is maintained over time the intervention is often seen by teachers as more effective (Baer et al., 1968, 1987) and would be perceived as more desirable given the potential for long-term positive outcomes for students such as higher academic achievement (Thomas et al. 2012). In addition, few studies included formal measures of the social validity of CBM apps on part of teachers and students. Studies that did assess social validity generally reported positive findings, with teachers perceiving I-Connect, SCORE IT, Tootling with ClassDojo and GBG with ClassDojo to be effective and useful when included as part of a CBM intervention. More research is needed to identify the conditions in which teachers might choose to use CBM apps, whether their use of apps promotes long-term and sustained behavior change for students, and whether apps are viewed as easy to use, effective, and preferred by both teachers and student.

As illustrated in the current review, CBM apps allow for the collection and sharing of data on students' behavior. As this information is sometimes stored on external servers, this raises concerns about how it is secured and who has access to it. The studies included in the present review did not provide a discussion of data storage and privacy considerations, although such considerations are important for teachers who use CBM apps. Future research is needed to explore ways to use data collection and sharing features within CBM apps to ensure that the privacy and confidentiality of students is protected.

Conclusion

The findings from this review suggest that CBM apps could play a positive role in supporting behavior change for students when used in the context of a well-designed intervention package. Of significance, the use of CBM apps were viewed favorably by both teachers and students. However, as decisions to use technology in the classroom are often based on perceptions of effectiveness rather than evidence (Sugar et al., 2004) additional research needs to be undertaken before a recommendation for their use in the classroom can be made with any degree of confidence. Teachers, school leaders, and policy makers should be mindful that CBM apps may not be suitable for all types of classrooms or student populations. They should carefully consider whether the use of CBM apps aligns with their schools' philosophy, mission and educational goals.

Author Contributions SM conceived of the review, conducted the literature search, reviewed all articles, undertook data analysis and drafted the manuscript. NF, EL, and BF assisted with determining the inclusion criteria, reviewed articles meeting the inclusion criteria, and helped draft the manuscript. EP also helped review articles meeting the inclusion criteria and draft the manuscript. All authors provided feedback on the manuscript drafts and read and approved the final manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

Data Availability The data that support the findings of this study are available from the corresponding author, SM, upon request.

Declarations

Conflicts of Interest The authors have no conflict of interest to declare that are relevant to the content of this article.

References

Alatupa, S., Pulkki-Råback, L., Hintsanen, M., Mullola, S., Lipsanen, J., & Keltikangas-Järvinen, L. (2011). Childhood disruptive behaviour and school performance across comprehensive school: A prospective cohort study. *Psychology, 2*(6), 542–551. https://doi.org/10.4236/psych.2011.26084

Apple Inc. (2007). App Store. In (Vol. iOS). https://www.apple.com/au/app-store/

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91–97. https://doi.org/10.1901/jaba.1968.1-91

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 20*(Winter), 313–327. https://doi.org/10.1901/jaba.1987.20-313

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*(2), 119–124. https://doi.org/10.1901/jaba.1969.2-119

Beaman, R., Wheldall, K., & Kemp, C. (2007). Recent research on troublesome classroom behaviour: A review. *Australasian Journal of Special Education*, *31*(1), 45–60. 10.1017%2Fs1030011200025586

Beckman, A., Mason, B. A., Wills, H. P., Garrison-Kane, L., & Huffman, J. (2019). Improving behavioral and academic outcomes for students with autism spectrum disorder: Testing an app-based self-monitoring intervention. *Education & Treatment of Children, 42*(2), 225–244. https://doi.org/10.1353/etc.2019.0011

Bedesem, P. L. (2012). Using cell phone technology for self-monitoring procedures in inclusive settings. *Journal of Special Education Technology, 27*(4), 33–46. https://doi.org/10.1177/016264341202700403

Bellack, A. S., & Hersen, M. (1998). *Comprehensive clinical psychology*. Pergamon.

Blackwell, C. K., Lauricella, A. R., Wartella, E., Robb, M., & Schomburg, R. (2013). Adoption and use of technology in early education: The interplay of extrinsic barriers and teacher attitudes. *Computers & Education, 69*, 310–319. https://doi.org/10.1016/j.compedu.2013.07.024

Bruhn, A. L., Rila, A., Mahatmya, D., Estrapala, S., & Hendrix, N. (2020). The effects of data-based, individualized interventions for behavior. *Journal of Emotional & Behavioral Disorders, 28*(1), 3–16. https://doi.org/10.1177/1063426618806279

Bruhn, A. L., Vogelgesang, K., Fernando, J., & Lugo, W. (2016). Using data to individualize a multicomponent, technology-based self-monitoring Intervention. *Journal of Special Education Technology, 31*(2), 64–76. https://doi.org/10.1177/0162643416650024

Bruhn, A. L., Vogelgesang, K., Schabilion, K., Waller, L., & Fernando, J. (2015). "I don't like being good!" Changing behavior with technology-based self-monitoring. *Journal of Special Education, 30*(3), 133–144. https://doi.org/10.1177/0162643415618911

Bruhn, A. L., Woods-Groves, S., Fernando, J., Choi, T., & Troughton, L. (2017). Evaluating technology-based self-monitoring as a tier 2 intervention across middle school settings. *Behavioral Disorders*, *42*(3), 119–131. 10.1177%2F0198742917691534

Bunch-Crump, K. R., & Lo, Y.-Y. (2017). An investigation of multitiered behavioral interventions on disruptive behavior and academic engagement of elementary students. *Journal of Positive Behavior Interventions, 19*(4), 216–227. https://doi.org/10.1177/1098300717696939

Cetin, H., & Cetin, I. (2018). Views of middle school students about ClassDojo education technology. *Acta Didactica Napocensia*, *11*(3–4), 89–96. https://doi.org/10.24193/adn.11.3-4.7.

Chaffee, R. K., Johnson, A. H., & Volpe, R. J. (2017). A meta-analysis of class-wide interventions for supporting student behavior. *School Psychology Review*, *46*(2), 149–164. https://doi.org/10.17105/spr-2017-0015.V46-2

Chen, W., Gu, X., & Wong, L.-H. (2017). To click or not to click: Effectiveness of rating classroom behaviors on academic achievement with tablets. *British Journal of Educational Technology, 50*(1), 440–455. https://doi.org/10.1111/bjet.12593

Chiarelli, M., Szabo, S., & Williams, S. (2015). Using ClassDojo to help with classroom management during guided reading. *Texas Journal of Literacy Education*, *3*(2), 81–88. https://files.eric.ed.gov/fulltext/EJ1110950.pdf

Christ, T. J., & Christ, J. A. (2006). Application of an interdependent group contingency mediated by an automated feedback device: An intervention across three high school classrooms. *School Psychology Review, 35*(1), 78–90. https://doi.org/10.1080/02796015.2006.12088003

ClassDojo Inc. (2020). *ClassDojo* [Mobile app]. In Apple App Store. http://www.apple.com/au/ios/app-store/

Clemons, L. L., Mason, B. A., Garrision-Kane, L., & Wills, H. P. (2016). Self-monitoring for high school students with disabilities: A cross-categorical investigation of I-Connect. *Journal of Positive Behavior Interventions, 18*(3), 145–155. https://doi.org/10.1177/1098300715596134

Clunies-Ross, P., Little, E., & Kienhuis, M. (2008). Self-reported and actual use of proactive and reactive classroom management strategies and their relationship with teacher stress and student behaviour. *Educational Psychology*, *28*(6), 693–710. 10.1080%2F01443410802206700

Crutchfield, S. A., Mason, R. A., Chambers, A., Wills, H. P., & Mason, B. A. (2015). Use of a self-monitoring application to reduce stereotypics behavior in adolescents with autism: A preliminary investigation of I-Connect. *Journal of Autism & Developmental Disorders, 45*(5), 1146–1155. https://doi.org/10.1007/s10803-014-2272-x

Dadakhodjaeva, K. (2017). *The Good Behavior Game: Effects on and maintenance of behavior in middle-school classrooms using ClassDojo* (Publication No. 363) [Doctoral dissertation, University of Southern Mississippi]. Dissertations.

DeJager, B., Houlihan, D., Filter, K. J., Mackie, P. F. E., & Klein, L. (2020). Comparing the effectiveness and ease of implementation of token economy, response cost, and a combination condition in rural elementary school classrooms. *Journal of Rural Mental Health, 44*(1), 39–50. https://doi.org/10.1037/rmh0000123

Dhir, A., Gahwaji, N. M., & Nyman, G. (2013). The role of the iPad in the hands of the learner. *Journal of Universal Computer Science*, *19*(5), 706–727. http://jucs.org/jucs_19_5/the_role_of_the/jucs_19_05_0706_0727_dhir.pdf

Dillon, M. B. M., Radley, K. C., Tingstrom, D. H., Dart, E. H., & Barry, C. T. (2019). The effects of Tootling via ClassDojo on student behavior in elementary classrooms. *School Psychology Review*, *48*(1), 18–30. https://doi.org/10.17105/SPR-2017-0090.V48-1

Elliott, C. (2017). *Increasing student perceptions of teacher caring using ClassDojo* [Masters theses, California State University]. Capstone Projects & Masters Theses.

Elswick, S., Casey, L. B., Zanskas, S., Black, T., & Schnell, R. (2016). Effective data collection modalities utilized in monitoring the good behavior game: Technology-based data collection versus hand collected data [Classroom Dynamics & Student Adjustment & Attitudes 3560]. *Computers in Human Behavior, 54*, 158–169. https://doi.org/10.1016/j.chb.2015.07.059

Evertson, C. M., & Weinstein, C. S. (2006). *Handbook of classroom management: Research, practice and contemporary issues*. Routledge.

Ford, W. B. (2018). *Evaluation of a positive version of the Good Behavior Game utilizing ClassDojo technology in secondary classrooms.* [Doctoral dissertation, University of Southern Mississippi].

Fox, R. A., Sharma, U., Leif, E. S., Stocker, K. L., & Moore, D. W. (2021). "Not Enough Time": Identifying Victorian teachers' perceptions of the facilitators and barriers to supporting improved student behaviour. *Australasian Journal of Special & Inclusive Education, 45*(2), 205–220. https://doi.org/10.1017/jsi.2021.6

Gettinger, M. (1988). Methods of proactive classroom management. *School Psychology Review, 17*(2), 227–242. https://doi.org/10.1080/02796015.1988.12085340

Great Schools Partnership. (2014). *The glossary of education reform.* https://www.edglossary.org/classroom-management/y. Accessed 26 Nov 2014

Greenwood, C. R., Delquadri, J., & Hall, R. V. (1984). Opportunity to respond and student academic performance. In W. T. Heward, T. D. Heron, D. Hill, & J. Trap-Porter (Eds.), *Behavior analysis in education* (pp. 58–88). University of Kansas, Juniper Gardens Children's Project.

Hammonds, L., Matherson, L. H., Wilson, E. K., & Wright, V. H. (2013). Gateway tools: Five tools to allow teachers to overcome barriers to technology integration. Special issue, Future of Education,*Delta Kappa Bulletin* (Fall), 36–40. https://www.proquest.com/openview/ff1969c450ea9e602338badb560ae70b/1?pq-origsite=gscholar&cbl=47978

Homer, R., Hew, K. F., & Tan, C. Y. (2018). Comparing digital badges-and-points with classroom token systems: Effects on elementary school ESL students' classroom behavior and English learning. *Educational Technology & Society*, *21*(1), 137–151. https://core.ac.uk/download/pdf/95558631.pdf

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Educational leadership*, *60*(8), 30–33. 10.1.1.182.106

Kaufman, J. S., Jaser, S. S., Vaughan, E. L., Reynolds, J. S., Di Donato, J., Bernard, S. N., & Hernandez-Brereton, M. (2010). Patterns in office referral data by grade, race/ethnicity, and gender. *Journal of Positive Behavior Interventions, 12*(1), 44–54. https://doi.org/10.1177/1098300708329710

Kinash, S., Brand, J., & Mathew, T. (2012). Challenging mobile learning discourse through research: Student perceptions of Blackboard Mobile Learn and iPads. *Australasian Journal of Educational Technology*, *28*(4), 639–655. https://doi.org/10.14742/ajet.832

King, J. D., Buolamwini, J., Cromwell, E. A., Panfel, A., Teferi, T., Zerihun, M., Melak, B., Watson, J., Tadesse, Z., Vienneau, D., Ngondi, J., Utzinger, J., Odermatt, P., & Emerson, P. M. (2013). A novel electronic data collection system for large-scale surveys of neglected tropical diseases. *PLoS One, 8*(9), e74570–e74570. https://doi.org/10.1371/journal.pone.0074570

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2014). *Effective classroom management strategies and classroom management programs for educational practice*. GION onderwijs/onderzoek.

Krach, S. K., McCreery, M. P., & Rimel, H. (2017). Examining teachers' behavioral management charts: A comparison of ClassDojo and paper-pencil methods. *Contemporary School Psychology, 21*(3), 267–275. https://doi.org/10.1007/s40688-016-0111-0

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*(2), 163–175. https://doi.org/10.1002/pits.20627

LeBel, T. J., Chafouleas, S. M., Britner, P. A., & Simonsen, B. (2013). Use of a daily report card in an intervention package involving home-school communication to reduce disruptive behavior in preschoolers. *Journal of Positive Behavior Interventions*, *15*(2), 103–112. 10.1177%2F1098300712440451

Lewis, R. (2008). Understanding pupil behaviour: Classroom management techniques for teachers. *David Fulton*. https://doi.org/10.4324/9780203883914

Lewis, T. J., Newcomer, L. L., Trussell, R., & Richter, M. (2013). Schoolwide positive behavior support: Building systems to develop and maintain appropriate social behavior. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 843–864). Routledge.

Lipscomb, A. H., Anderson, M., & Gadke, D. L. (2018). Comparing the effects of ClassDojo with and without Tootling intervention in a postsecondary special education classroom setting. *Psychology in the Schools, 55*(10), 1287–1301. https://doi.org/10.1002/pits.22185

Lizzy B. Good Behavior Consulting. (2015). SCORE IT [iPad app]. In (Vol. Apple iPad).

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin*, *90*(1), 1–94. https://doi.org/10.21236/ada086584

Lunenburg, F. C. (2011). Goal-setting theory of motivation. *International Journal of Management, Business & Administration*, *15*(1), 1–6. http://www.nationalforum.com/Electronic%20Journal%20Volumes/Lunenburg,%20Fred%20C.%20Goal-Setting%20Theoryof%20Motivation%20IJMBA%20V15%20N1%202011.pdf

Lynne, S., Radley, K. C., Dart, E. H., Tingstrom, D. H., Barry, C. T., & Lum, J. D. K. (2017). Use of a technology-enhanced version of the good behavior game in an elementary school setting. *Psychology in the Schools, 54*(9), 1049–1063. https://doi.org/10.1002/pits.22043

Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students with challenging behavior. *Journal of School Psychology, 50*(5), 625–654. https://doi.org/10.1016/j.jsp.2012.06.001

Marzano, R., & Marzano, J. S. (2003). The key to classroom management. *Educational leadership*, *61*(1), 6–13. https://www.researchgate.net/profile/Robert-Marzano/publication/283749466_The_Key_to_Classroom_Management/links/56f26c0908aed354e57293d3/The-Key-to-Classroom-Management.pdf

McIntosh, K., & Goodman, S. (2016). *Integrated multi-tiered systems of support: Blending RTI and PBIS*. Guilford Press.

Murphy, M., Mathis, E., Weaver, A. D., & Dart, E. (2019). Using antecedent-based strategies to address motivation in behavioral interventions. *Communique (National Association of School Psychologists)*, *48*(2), 1–23. https://go.gale.com/ps/i.do?p=AONE&u=monash&id=GALE%7CA602824782&v=2.1&it=r

O'Brien, C. & Aguinaga, N. (2014). *Using Class Dojo with an interactive whiteboard and ipad projection to promote positive behavioral interventions and supports at the classroom level*. In M. Searson & M. Ochoa (Eds.), *Proceedings of SITE 2014--Society for Information Technology & Teacher Education International Conference. Jacksonville*, FL. Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/primary/p/131249/.

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*(4), 357–367. https://doi.org/10.1016/j.beth.2008.10.006

Parsonson, B. S. (2012). Evidence-based classroom behaviour management strategies. *Kairaranga*, *13*(1), 16–23. https://files.eric.ed.gov/fulltext/EJ976654.pdf

Pokorski, E. A., Barton, E. E., & Ledford, J. R. (2019). Assessing the differential effects of known and mystery rewards in a preschool-based group contingency. *Journal of Early Intervention, 41*(3), 256–275. https://doi.org/10.1177/1053815119841088

Pustejovsky, J. E., Chen, M., Grekov, P., & Swan, D. M. (2023). Single-case effect size calculator (Version 0.7.2.9999) [Computer software]. https://jepusto.shinyapps.io/SCD-effect-sizes/

R Core Team. (2023). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Radley, K. C., Dart, E. H., & O'Handley, R. D. (2016). The Quiet Classroom Game: A class-wide intervention to increase academic engagement and reduce disruptive behavior. *School Psychology Review*, *45*(1), 93–108. https://doi.org/10.17105/SPR45-1.93-108

Ratcliff, N. J., Jones, C. R., Costner, R. H., Savage-Davis, E., & Hunt, G. H. (2010). The elephant in the classroom: The impact of misbehavior on classroom climate. *Education*, *131*(2), 306–314. https://eds.s.ebscohost.com/eds/detail/detail?vid=0&sid=fef529ae-539e-4101-b4cb-970b92896a95%40redis&bdata=JkF1dGhUeXBlPXNoaWImc2l0ZT1lZHMtbGl2ZQ%3d%3d#AN=59347948&db=trh

Riden, B. S., Markelz, A. M., & Randolph, K. M. (2019). Creating positive classroom environments with electronic behavior management programs. *Journal of Special Education Technology, 34*(2), 133–141. https://doi.org/10.1177/0162643418801815

Robacker, C. M., Rivera, C. J., & Warren, S. H. (2016). A token economy made easy through ClassDojo. *Intervention in School & Clinic, 52*(1), 39–43. https://doi.org/10.1177/1053451216630279

Robichaux, N. M., & Gresham, F. M. (2014). Differential effects of the mystery motivator intervention using student-selected and mystery rewards. *School Psychology Review, 43*(3), 286–298. https://doi.org/10.1080/02796015.2014.12087428

Rosenbloom, R., Mason, R. A., Wills, H. P., & Mason, B. A. (2016). Technology delivered self-monitoring application to promote successful inclusion of an elementary student with autism. *Assistive Technology, 28*(1), 9–16. https://doi.org/10.1080/10400435.2015.1059384

Saeger, A. M. (2017). *Using ClassDojo to promote positive behaviors and decrease undesired behaviors in the classroom* [Masters dissertation, Rowan University]. Rowan Digital Works: Theses & Dissertations.

Shinn, M. R., Ramsey, E., Walker, H. M., Stieber, S., & O'Neill, R. E. (1987). Antisocial behavior in school settings: Initial differences in an at risk and normal population. *Journal of Special Education, 21*(2), 69–84. https://doi.org/10.1177/002246698702100207

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education & Treatment of Children, 31*(3), 351–380. https://doi.org/10.1353/etc.0.0007

Skinner, C. H., Cashwell, T. H., & Skinner, A. L. (2000). Increasing tootling: The effects of a peer-monitored group contingency program on students' reports of peers' prosocial behaviors. *Psychology in the Schools, 37*(3), 263–270. https://doi.org/10.1002/(SICI)1520-6807(200005)37:3%3c263::AID-PITS6%3e3.0.CO2-C

Sugar, W., Crawley, F., & Fine, B. (2004). Examining teachers' decisions to adopt new technology. *Educational Technology & Society*, *7*(4), 201–213. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=862d0a4c870323a7688e375cf38ad9675caf0116

Sullivan, A., Johnson, B., Conway, R., Owens, L., & Taddeo, C. (2014). Punish them or engage them? Teachers' views of unproductive student behaviours in the classroom. *Australian Journal of Teacher Education*, *39*(6), 43–56. https://doi.org/10.14221/ajte.2014v39n6.6

Szwed, K., & Bouck, E. C. (2013). Clicking away: Repurposing student response systems to lessen off-task behavior. *Journal of Special Education Technology, 28*(2), 1–12. https://doi.org/10.1177/016264341302800201

Thomas, C. N., Hassaram, B., Rieth, H. J., Raghavan, N. S., Kinzer, C. K., & Mulloy, A. M. (2012). The Integrated Curriculum Project: Teacher change and student outcomes within a university-school professional development collaboration. *Psychology in the Schools, 49*(5), 444–464. https://doi.org/10.1002/pits.21612

Vogelgesang, K. L., Bruhn, A. L., Coghill-Behrends, W. L., Kern, A. M., & Troughton, L. C. W. (2016). A single-subject study of a technology-based self-monitoring intervention. *Journal of Behavioral Education, 25*(4), 478–497. https://doi.org/10.1007/s10864-016-9253-4

Walter, H. J., Gouze, K., & Lim, K. G. (2006). Teachers' beliefs about mental health needs in inner city elementary schools. *Journal American Academy Child Adolescent Psychiatry, 45*(1), 61–68. https://doi.org/10.1097/01.chi.0000187243.17824.6c

What Works Clearinghouse. (2022). *What Works Clearinghouse: Procedures and standards handbook, Version 5.0*. U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Handbooks

Wills, H. P., & Mason, B. A. (2014). Implementation of a self-monitoring application to improve on-task behavior: A high-school pilot study. *Journal of Behavioral Education, 23*(4), 421–434. https://doi.org/10.1007/s10864-014-9204-x

Wilson, R. M. (2017). *ClassDojo.com: The effects of a digital classroom management program* (Publication No. 10257246) [Doctoral dissertation, Trevecca Nazarene University]. ProQuest Dissertations.

Wolf, P. J. (2015). *ClassDojo: An awesome progress monitoring tool!* Georgia Southern University.