Self-Management Strategies and Academic Engagement for Autistic Students: A Multilevel Meta-Analysis of Single-Case Experimental Designs

Jamie Gahtan[1] Mariola Moeyaert[1] Kristie Asaro-Saddler[1] Elie ChingYen Yu[1] Edmund Orlowski[1]

[1] University at Albany, State University of New York

Department of Educational and Counseling Psychology

Author Contact:
Jamie Gahtan, MSEd https://orcid.org/0000-0002-0847-4580; Email: jgahtan2@albany.edu
Mariola Moeyaert, PhD https://orcid.org/0000-0003-1453-8162; Email: mmoeyaert@albany.edu
Kristie Asaro-Saddler, PhD https://orcid.org/0000-0003-2247-7419; Email: ksaddler@albany.edu
Elie ChingYen Yu, MS https://orcid.org/0000-0002-6716-3643; Email: cyu5@albany.edu
Edmund W. Orlowski, MS https://orcid.org/0000-0002-0696-5693; Email: eorlowski@albany.edu

Funding statement: This review received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interests: The authors declare none.

Short title: Self-Management and Academic Engagement

* This paper has undergone peer review and has been accepted for publication in the Journal of Positive Behavioral Interventions. This version is the accepted manuscript, prior to final proof and formatting.

Abstract

This single-case design meta-analysis examined the effects of self-management strategies (SM) on academic engagement in autistic students. Prior meta-analyses reported positive outcomes but did not account for the nested structure of single-case designs. Using multilevel modeling, we addressed this limitation and examined moderators of implementation practices. Eligible studies were peer-reviewed, published in English, school-based, included at least one autistic participant, used a single-case design with a baseline and intervention phase, employed time intervals for SM, and targeted positive outcomes. We searched ProQuest, EBSCO, and Web of Science through March 10, 2024. Risk of bias was assessed using the SCD RoB tool. 14 studies (29 participants) met inclusion criteria; SM increased academic engagement by 49.87% (SE = 4.33, 95% CI [39.76, 59.59], $p < .001$). Variability existed across participants and studies, with 9 studies (17 participants) contributing to moderator analyses. Classroom subject, classroom type, technology-based prompting, and the time interval of SM did not significantly explain variability. Limitations include small sample sizes, exclusion of three studies due to missing information, and high risk of bias in assessors' knowledge of study phase. Overall, SM has the potential of being a flexible strategy supporting autistic students' academic engagement.

Keywords: Autism Spectrum Disorder, self-management strategies, evidence-based practices, single-case designs

Introduction

Self-management (SM) strategies are evidence based techniques that help students achieve desired behaviors by observing, assessing, and modifying their own behavior (Cooper et al., 2020; Schulze, 2016). SM strategies are an effective practice for many students, including autistic students (Carr et al., 2014; Lee et al., 2007; Scheibel et al., 2024). Autism is described as a developmental disability categorized by differences in communication and social interactions, along with repetitive behaviors, restricted interests, and atypical sensory processing (American Psychiatric Association, 2022). Severity ranges from Level 1, requiring support, to Level 3, requiring substantial support, and is often accompanied by differences in language use and co-occurring intellectual disabilities. As a result, some autistic students may exhibit behaviors that interfere with learning and academic engagement (Sparapani et al., 2016). Many of the challenges autistic students face that impact academic performance often stems from factors outside of academic ability (Adams et al., 2025), such as non-academic skills including self-regulation, sustained attention, and engagement that impact classroom performance (Carnahan et al., 2009; Morningstar et al., 2017).

Academic engagement, defined as the level of participation and investment in educational activities, is a critical component of academic success (Bangert-Drowns & Pyke, 2001; Lawson & Masyn, 2015). However, research has shown that autistic students tend to demonstrate lower engagement and spend less time on task during instruction compared to their typically developing peers and students with other disabilities (Carnahan et al., 2009). Given this, identifying strategies that promote engagement and on-task behavior is essential for supporting the academic outcomes of autistic students.

SM interventions, which have been used to support engagement and on-task behavior, are particularly well-suited for autistic students because they can be individualized to reflect the unique profiles of students and promote autonomy in the classroom (Briesch et al., 2018; Odom et al., 2012). Several meta-analyses of single-case design research have reported positive effects of SM intervention on combined outcomes, including academic, social, behavioral, adaptive, and communication skills, in autistic students (Carr et al., 2014; Lee et al., 2007; Scheibel et al., 2024). However, the term SM has been applied to a wide variety of intervention approaches that differ in key implementation features, such as classroom setting, grade level, self-reporting format, and frequency of student self-recording (Briesch et al., 2018; Scheibel et al., 2024). While prior reviews have acknowledged this variability and described the proportion of studies using specific implementation characteristics, they have not systematically examined whether and how these features independently are related to academic outcomes (Briesch et al., 2018; Scheibel et al., 2024). This gap is especially true for academic engagement, which is a consistently used outcome in SM research (Scheibel et al., 2024). Instead, conclusions of effectiveness across implementation features have been based on the distribution of different implementation characteristics across studies or the effects of overall configurations of SM elements.

Moreover, previous synthesis often relied on analytic techniques that do not account for the nested nature of single case design data (Moeyaert et al., 2018, 2023). This study aims to build on the existing evidence base by applying a multilevel meta-analytic approach to explore the effectiveness of SM interventions for autistic students' academic engagement, as well as on additional academic outcomes when sample size permits. This approach integrates statistical advances that consider the nested nature of the data and allow for the systematic investigation of moderators (Moeyaert et al., 2018, 2023). By empirically investigating implementation characteristics across studies, this review aims to advance the field's understanding of how variations in SM intervention characteristics may relate to academic engagement for autistic students.

Academic Performance and Engagement in Autistic Students

Academic achievement among autistic students varies widely (Estes et al., 2011; Kim et al., 2018), with many autistic students performing below the threshold of anticipated academic abilities across a variety of subject areas and domains (Keen et al., 2023). Student engagement, in particular, is widely recognized as a key contributor to learning and academic achievement (Bangert-Drowns & Pyke, 2001; Lawson & Masyn, 2015). Although engagement is multidimensional—including affective engagement, behavioral engagement, and cognitive engagement (Lawson & Lawson, 2013; Lawson & Masyn, 2015)—in classroom settings, engagement is often operationalized unidimensionally as observable behavior such as looking at the teacher, remaining seated, and proportion of time working on the appropriate task (Carnahan et al., 2009; Cihak et al., 2010; Xu et al., 2017). Autistic students often exhibit low levels of behavioral engagement during instruction, which can reduce their availability for learning (Carnahan et al., 2009; Sparapani et al., 2016).

In addition, autistic students may experience additional academic challenges with learning, attention, and processing speed compared to their typically developing peers (Mayes & Calhoun, 2007), and such skills as inhibitory control and attention skills has been found to be associated with task engagement and academic skills in typically developing students (e.g., emergent literacy and language skills; Bohlmann & Downer, 2016). As a result, autistic students may be less available for learning by spending less time on task—behavior that demonstrates a student is behaviorally engaged in the appropriate learning activity (Carnahan et al., 2009)—

which can adversely impact their academic performance. Therefore, it is not surprising that a majority of studies investigating the effectiveness of SM for academics have focused on the reporting of on-task behavior as an outcome measure (see Briesch et al., 2018 for a review). Furthermore, because on-task behavior is a component of behavioral engagement (Fredricks, 2022), and they often have overlap in their defined measurement (e.g., both might contain looking at the teacher; Callahan et al., 1999; Xu et al., 2017), a previous review has grouped together author reported outcomes of engagement and on-task behavior as one outcome: engagement (Scheibel et al., 2024).

Self-Management

One way to support autistic students in the classroom is through evidence-based practices (EBPs), which are demonstrated through scientifically rigorous research to have positive outcomes. Two independent organizations, the National Professional Development Center on ASD (Hume et al., 2021) and the National Autism Center (2015), have developed a list of evidence-based practices and programs for autistic people. Among these practices is SM. The National Clearinghouse on Autism Evidence and Practice describes SM as distinguishing between appropriate and inappropriate behaviors, planning, executing, and monitoring successful behaviors, and rewarding themselves when appropriate (Steinbrenner et al., 2020). Though often equated with self-monitoring, SM is a broader construct that encompasses a range of strategies, including goal setting, self-evaluation and self-reinforcement (Chia et al., 2018; Rafferty & Asaro-Saddler, 2021). SM strategies can be easily adaptable or individualized, which may be ideal for autistic students (Scheibel et al., 2024) given their varying nature (American Psychiatric Association, 2022; Kim et al., 2018).

Research has documented substantial variability in how SM interventions are implemented. In a broader systematic review examining characteristics of SM interventions in 56 studies across various populations and outcomes, Briesch, Daniels, and Beneville (2019) found large variability in how SM interventions are implemented. Some of the common features of reporting included differences in the frequency of self-reporting, as well as variations in implementation settings (e.g., general education vs. special education classrooms) and grade levels. Although they found strong effects for SM interventions overall, the high degree of variation across studies limited the ability to isolate and evaluate the impact of specific features. They were able to calculate effect sizes only for intervention configurations that appeared in at least three studies with just one varying feature, for example, the type of diagnosis for elementary students in special education classes using prompt-based SM, the effect of graphing with SM, or the effect of using checklists for SM procedures. This constrained efforts to compare effect sizes of many of the specific implementation features (e.g., setting and frequency of self-reporting). Moreover, SM interventions have also occurred during a variety of class subjects (e.g., math, language arts, science; Cihak et al. 2010; Holifield et al. 2010) and have used a variety of materials for self-reporting including technology-based applications or paper and pens (Chia et al., 2018; Scheibel et al., 2024).

Despite variability in implementation, SM has several practical benefits in classroom settings. SM promotes independence, can be applied in diverse environments, and often supports autistic students with generalizing behavioral changes, which is a common challenge for many autistic individuals (McDougall 1998; Rafferty & Asaro-Saddler, 2021). In the classroom setting, SM allows for minimal management by teachers and, in turn, can maximize instructional time (Cooper et al., 2020; Southall & Gast, 2011). Moreover, general education teachers have reported a preference for using SM in integrated classroom settings as they require minimal facilitation and implementation actions by teachers (Southall & Gast, 2011). This finding is particularly relevant as the number of autistic students placed full-time in general education classrooms has increased (National Center for Education Statistics, 2023).

Previous meta-analyses of single-case design studies have evaluated the effectiveness of SM for autistic people (e.g., Carr et al., 2014; Chia et al., 2018; Scheibel et al., 2024). Carr and colleagues (2014) analyzed 23 single-case design studies using the What Works Clearinghouse standards (Kratochwill et al., 2010) to identify quality, which includes judgments, for example, on the number of participants and demonstrations of effect in each study. They also used the percentage of non-overlapping data (PND)—what percentage of observations in the treatment phase do not overlap with the baseline phase (Scruggs et al., 1987)—to quantify the intervention effect. Results indicate that SM interventions effectively increased academic skills for students of all ages and levels of ability and that generalization and maintenance rates were high, with a mean PND of 84.3% (Carr et al., 2014). Chia and colleagues (2018) also examined 12 single-case studies using the What Works Clearinghouse standards, focusing specifically on studies using technology to support SM interventions. Using the percentage of non-overlapping data, they found the interventions to be effective, with a mean PND = 78.8% and a median PND = 88.9%. Most recently, Scheibel et al. (2024) conducted a meta-analysis across multiple domains (e.g., social, behavioral, academic) and reported that studies were of varying quality, with just over a quarter of studies (26%) being excluded due to low quality or rigor, as analyzed using the Single Case Analysis and Review Framework 2.0 (SCARF; Ledford et al., 2020). Of the reviewed studies, Scheibel and colleagues (2024) focused on 24 treatment effects on the outcome of engagement across five studies. Using the Log Response Ratio (LRR), or the log of the proportion of the mean of the intervention data and the mean of the baseline data (Pustejovsky et al., 2018), they determined that using SM strategies positively increased engagement behavior (LRR%=166%, Scheibel et al., 2024).

Despite promising findings, previous meta-analyses face methodological limitations. First, PND and LRR were used to compute effect sizes. Although both methods allow for a combined understanding of effects across single case design studies, they have some limitations. PND, for instance, is at risk for outliers (Rakap, 2015) and can fail to detect meaningful effects when there are observable visual changes in performance level, trend, or variability between phases in graphical data displays (Lenz, 2013). The LRR index can be heavily influenced by procedural characteristics of a study's design (Pustejovsky et al., 2018; 2019). As there is much diversity in single-case design methodologies, nuanced procedural aspects can potentially introduce bias and affect the reliability of results.

Furthermore, both PND and LRR are case-specific effect sizes and are traditionally aggregated across cases and across studies by using a simple average (Pustejovsky et al., 2018; Rakap, 2015). Consequently, the nested nature of single case design data is ignored, wherein cases are nested within participants and participants are nested within studies (Moeyaert et al., 2018). Therefore, case-level (e.g., classroom subject) and study-level (e.g., intervention modality) moderators cannot be investigated. For this reason, a multilevel meta-analysis regression model is recommended and applied in the current study, as this advanced technique allows for estimating effect sizes while considering the nested nature of the data (Moeyaert, 2014; Moeyaert et al., 2018; 2023). An overall average weighted effect size across cases and studies is obtained in addition to an estimate of between-case and between-study variance in intervention effectiveness. Moreover, case and study-level moderators can be added to the model in an attempt to explain intervention heterogeneity by investigating the effect for each case, between cases in a study, and between multiple studies (Moeyaert, 2014; Moeyaert et al., 2018).

Current Study

In this study, we aimed to explore the effects of SM on academic engagement and academic outcomes more broadly for autistic students. Specifically, we used a three-level hierarchical meta-analysis to estimate the effects of SM on academic engagement and investigate variability in intervention effectiveness between studies and participants. This extends previous work by allowing for a more comprehensive understanding of the nested nature of data while facilitating the inclusion of other implementation factors at the case and study level, such as classroom type, academic subject, method of students' SM self-reporting, and time intervals of student self-reporting. We also include additional databases from those used in previous meta-analyses (Carr et al., 2014; Scheibel et al., 2024) to expand our search and include dissertations in our search criteria. Our goal was to produce a more comprehensive understanding of the effect of SM on academic engagement and identify key factors that contribute to the effectiveness of SM.

Based on previous meta-analyses that have discovered positive academic effects of SM for autistic students (e.g., Carr et al., 2014; Scheibel et al., 2024), we hypothesized that SM will positively affect academic outcomes, particularly academic engagement, for autistic students. Additionally, as many studies discuss the ability to adapt and create individualized plans of SM strategies (Cooper et al., 2020; Rafferty & Asaro-Saddler, 2021), we expected that the moderators of classroom type (i.e., self-contained or general education) and subject (i.e., science and math or other) will account for little variance in the effect of the SM intervention. Additionally, we anticipated that autistic students would broadly vary in their initial performance (i.e., mean baseline) before the SM implementation and intervention effectiveness of academic engagement (i.e., treatment effect; e.g., Chen et al., 2018; Kim et al., 2018). As previous meta-analyses have not investigated the frequency of student self-monitoring and the use of applications vs non-applications for self-monitoring, these two moderators are exploratory.

Methods

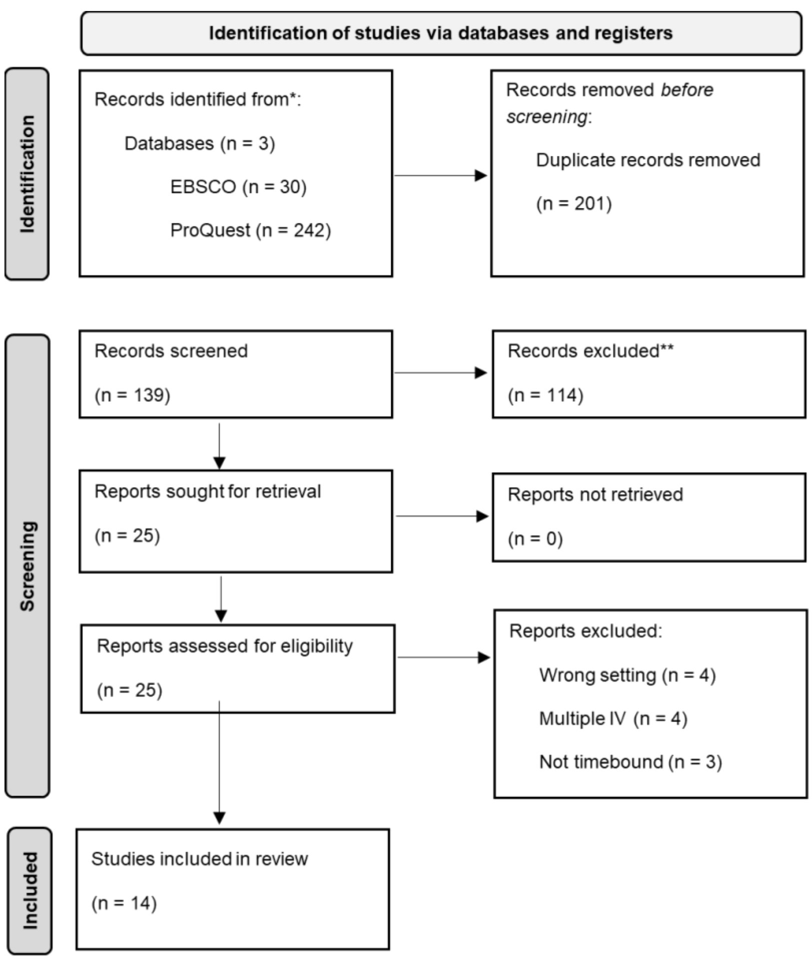
Literature Search

A systematic review was conducted to identify eligible studies investigating the impact of SM on autistic students' academic engagement and academic outcomes more broadly. To expand the previous meta-analysis, we consulted the databases APA PsycInfo, Education Resources Information Center (ERIC), and ProQuest Dissertation and Theses Global within ProQuest, as these databases were previously used by Carr et al. (2014) and Scheibel et al. (2024) and added the databases ProQuest (APA PsycArticles and Medline), EBSCO (Academic Search Complete, Education Source, Psychology and Behavioral Sciences Collection, Teacher Reference Center), and Web of Science.

Search strings were adapted from terms used in Carr et al. (2014). To further isolate relevant results, we used ChatGPT (OpenAI., 2023) following script prompts from Wang et al. (2023) to develop appropriate search strings. The following four keyword strings isolated single case design studies that use SM strategies to target academic outcomes for autistic students: *(1) self-monitoring* OR *self-management* OR *self-recording* OR *self-reinforcement* OR *self-evaluation* OR *self-regulation* OR *R self-regulate* with (2) autism** OR *autistic** OR *ASD* OR *autism spectrum disorder** OR *persuasive developmental disorder** OR *PDD** OR *Asperger* with 3) single case* OR *single subject* OR *N of 1 Or small N* OR *multiple baseline design* or reversal design* OR withdrawal design* OR *ABAB* OR *ABA or alternative treatment* design* OR *alternating treatment* with 4) academic** OR *writing* OR *reading* OR *literacy* OR *language arts*

OR *math* OR *science* OR *social studies* OR *history* OR *on-task**. Additional search criteria included peer-reviewed articles in English published before March 10th, 2024. Boolean search strings were applied to title and abstracts. Three hundred thirty-six articles were retrieved across the databases (see Figure 1). For reliability purposes, an independent search using the same procedure was performed by the fourth author, which yielded identical results. An ancestorial search also occurred from meta-analyses previously published in the past ten years (Carr et al., 2014; Chia et al., 2018; Scheibel et al., 2024). Four additional articles were identified for screening.

Figure 1
Flow Process of Eligible Studies



The software program Rayyan assisted with deduplication and tracking the screening process (Ouzzani et al., 2016). Using the deduplication tool, the first author reviewed the 201 articles flagged as possible duplicates. When a dissertation was published as an article, the version with more information was maintained. After careful consideration and removal of identical articles, a total of 139 unique articles were retained.

To be considered for inclusion in this meta-analysis, each article initially underwent an abstract and title screening, isolating articles that met the following five criteria. First, for a study to be eligible, an autism diagnosis or priorly used terminology for autism, such as Asperger's and pervasive developmental disorder not otherwise specified (PDD-NOS; Oberman & Kaufmann, 2020), needed to be present for at least one participant, and only those participants were included in the analysis. Second, the study needed to take place in a school-based setting, such as a preschool, elementary school, middle school, high school, or post-secondary setting. Third, studies needed to employ single-case designs that included both a baseline and intervention phase, such as multiple baseline designs (MBD) and reversal designs (ABAB). Fourth, the intervention was an independent SM where students self-document academic performance (e.g., on task, engagement, task completion) at set time intervals that reflect the outcome measured. Studies with students who only self-monitored at the conclusion of the entire documented session were excluded. Fifth, the study must have focused on increasing positive academic outcomes (e.g., increasing on-task behavior or engagement; c.f., decreasing off-task behavior). Screening questions were created by the first author and are available at https://osf.io/947ec/?view_only=9df6b9aaccbf4490ad4289faef07318f. Studies with unclear inclusion criteria during the abstract and title screening were retained for full-text screening.

The first author trained the fourth author on screening procedures. The training began with a review and discussion of the screening questions. The first author then demonstrated the screening process by evaluating two articles in accordance with the screening questions, while thinking aloud to explain each step. Following this, the authors screened two articles together. Next, two articles were screened independently and compared for adequate reliability before continuing independently. The fourth author independently screened the abstracts and titles of 20.14% of the 139 articles. The first and fourth authors were aligned on 25 out of 28 (89.28%) of abstract and title screening decisions. Discussions were used to resolve all discrepancies. A total of 25 articles were retained for full-text review.

Figure 1 depicts the article screening process. Of the 25 articles that underwent full-text screening, 14 met the inclusion criteria. For reliability of full-text screening, the fourth author also independently screened 20% of eligible studies. The first and fourth authors were 100% aligned on all full-text screening decisions (5 out of 5 decisions). Articles were excluded due to SM occurring outside of a school setting (N =4), having multiple intervention techniques (N =4), and not incorporating timebound reporting (N = 3).

Data Extraction and Inter-rater Agreement

Raw data were obtained from studies graphs (Beckman et al., 2019; Callahan et al., 1999; Cihak et al., 2010; Clemons et al., 2016; Finn et al., 2015; Holifield et al., 2010; Huffman et al., 2019; Legge et al., 2010; Romans et al., 2020; Rosenbloom et al., 2017; Rosenbloom et al., 2016; Siko, 2018; Stasolla et al., 2014; Xu et al., 2017). Data extraction occurred using the software program WebPlotDigitalizer (Rohatgi, 2011), a recommended open-source, user-friendly data extraction tool, following the procedures in Moeyaert et al. (2016). Nine hundred sixteen data points were extracted from 20 figures amongst the 14 studies.

Lin's concordance coefficient was used to determine inter-observation agreement (IOA) for raw extracted data points (Lin, 1989). Lin's concordance coefficient is an appropriate estimate of agreement for continuous variables and is recommended for test-rest of raw data retrieval from single case design graphs (Moeyaert et al., 2016). The fifth author independently extracted data points from all included studies (n = 14; 916 raw data points), and the first author independently extracted data from four studies (199 raw data points). Both coders extracted 21.72% of raw data. The IOA for extracted data was 98.30%.

Study Characteristics Coding

At the study level, we coded the number of autistic participants, the country in which the intervention was conducted, the predominant grade level of participants (i.e., Preschool/PreK, Elementary [K–5], Middle School [6–8], High School [9–12], or Postsecondary), class size, intervention setting (i.e., general education, special education/self-contained, or inclusive co-taught classrooms), school type (public vs. private), the individual delivering the intervention or training (i.e., general education teacher, special education teacher, peer, teaching assistant, researcher, or other), the length of SM training, the recording interval for student self-management in seconds, the intervention modality (e.g., I-Connect, audio prompting), the overall duration of the intervention, the frequency of observer data collection, and the outcome variables assessed. At the participant level, we coded each participant's name, age, gender, grade level, diagnosis, and instructional subject during the SM intervention.

The first author created a coding form and modeled the coding process for the fourth author using one of the included studies. The first and fourth authors then jointly coded one study together. Following this, the fourth author independently coded study characteristics for three of the fourteen included studies, while the first author coded all fourteen studies. Any discrepancies were resolved through discussion. The IOA for study characteristics was 86.96%.

Risk of Bias in Individual Studies

The SCD RoB tool was used to understand the risk of bias among studies (Reichow et al., 2018). The SCD RoB tool was developed based on Cochrane risk of bias criteria and adapted to consider indicators of single case designs and standards appropriately. Using this tool, each potential domain of bias receives a risk of bias rating of having a low risk of bias, an unclear risk of bias, or a high risk of bias for each study (see Reichow et al., 2018 for a discussion on each type of bias). For the IOA of risk of bias coding, two authors coded 3 of the studies using the RoB tool. The IOA for risk of bias was 88.46%.

Plan for Analysis

The multilevel meta-analysis framework, which considers the nested data structure, was used to combine the regression-based effect sizes across cases and across studies (Moeyaert et al., 2014). Moreover, multilevel analysis captures the variability in an intervention's effectiveness between studies and between cases and allows modeling moderators to explain variability (Moeyaert et al., 2023).

Amongst the 14 studies, 29 participants had an academic outcome described by authors as engagement or on-task behavior. There was much overlap in how studies conceptualized engagement versus on-task. For instance, Cihak et al. (2010, p. 139) defined task engagement as “(a) being in one’s seat, (b) looking at the materials or teacher as requested, (c) writing letters or words related to the assigned task, and d) complying with instructions with 4 s,” while Stasolla et al. (2014, p. 473) defined on task as the student “(a) remained sitting on their desk, (b) kept silent, (c) listened to the teacher’s explanation, (d) focused on their task, gazing on their sheet, reading carefully the task and achieving it." Considering that studies conceptualizations of engagement and on-task behavior similarly intersect and are all encompassed within the behavioral component of academic engagement (Fredricks, 2022), we aggregated studies that used either outcome into a single category of academic engagement for analysis.

Other academic outcomes identified within the screening procedures included accuracy (Beckman et al., 2019; Holifield et al., 2010; Romans et al., 2020) and activity completion (Rosenbloom, 2017). However, because there were only ten individual participants across the four studies that included accuracy or activity completion, these outcomes were excluded from this meta-analysis. Accordingly, all analysis in this study focuses on academic engagement.

The statistical software R (Version 4.4.0) was used to run all analyses. Package *lme 4* (Bates et al., 2015) was used to run the random effects meta-analysis and meta-regression (i.e., random effects meta-analysis with the inclusion of moderators).

Two models were run to answer our research questions. Our first model (Model 1) focused on the treatment effect (i.e., the change between the mean baseline level and mean intervention level). The second model (Model 2) builds on the first model, exploring treatment effects while including moderators to understand the role of the characteristics of SM implementation in the effectiveness of SM.

The following equation represents Model 1:

$$Y_{ijk} = \gamma_{000} + v_{00k} + u_{0jk} + (\gamma_{100} + v_{10k} + u_{1jk})\, D_{ijk} + e_{ijk} \tag{1}$$

In this equation, academic engagement is referred to as Y_{ijk} obtained at the specific observation (*i*= 0, 1,…*I*), within each case *j* (*j*=1, 2,…*J*), and encompassed within the study *k* (*k*= 1, 2, 3, 4, ...K). D_{ijk} represents the study phase and is dummy coded with zero for the baseline phase and one for the intervention phase. The expected baseline level at the beginning of the baseline phase across cases and studies is represented by γ_{000}, while γ_{100} represents the expected change in level between the baseline phase and the intervention phase. The deviation of study k from the overall expected baseline is captured by v_{00k}. v_{10k} stands for the deviation of study k from the overall expected change in level. The values u_{0jk} and u_{1jk} represent the deviations of individual cases *j* from study k from the overall expected baseline and overall expected change from baseline to treatment phase (see Moeyaert et al., 2014 for more information).

The second multilevel analysis (Model 2) was conducted to explain variability in intervention effectiveness between participants and between studies. Model 2 includes four moderators: (1) academic subject (dummy codded: math/science or language arts/social studies/communication), (2) classroom type (dummy codded: special education or general/inclusive), (3) modality of students self-monitoring (dummy codded: application vs other modality), and (4) time interval in seconds of student self-reporting (continuous variable), were considered.

The following equation represents Model 2:

$$y_{ijk} = \gamma_{000} + v_{00k} + u_{0jk} + \left(\gamma_{100} + \gamma_{110} Subject_{11k} + \gamma_{101} Classroom_{101} + \gamma_{102} Modality_{101} + \gamma_{103} TimeInterval_{101} + v_{10k} + u_{1jk}\right) D_{ijk} + e_{ijk} \quad (2)$$

$$with\ e_{ijk} \sim N(0, \sigma_e^2), [u_{0jk}\ u_{1jk}] \sim N(0, \Sigma_u), and\ [v_{00k}\ v_{10k}] \sim N(0, \Sigma_v)$$

This model expands on the previous model to include moderators. The parameters in Model 2 that are also present in Model 1 can be interpreted in the same manner described in Model 1. The moderator academic subject ($\gamma_{110} Subject_{11k}$) is at the participant level. Classroom type ($\gamma_{101} Classroom_{101}$), modality of students self-monitoring ($\gamma_{102} Modality_{101}$) and time interval of students self-reporting ($\gamma_{103} TimeInterval_{101}$) are moderators at the study level.

Results

A total of 14 studies met all five inclusion criteria and were included in the multilevel meta-analysis. There was a total of 29 autistic participants (*M* = 2.7; *SD* = 1.07) across studies. Participant ages ranged from 7 to 22 years ($M_{age} = 12.45$ years; *SD* = 4.30). Each study was coded for various study-level and participant-level characteristics, which were not included in the current analysis. See Tables 1 and 2 for additional study and participant characteristics, respectively. Furthermore, additional study and participant information is available on OSF.

Table 1

Summary of Study Characteristics

Study	Academic Behavior	Intervention Modality	School Type	Classroom Type	Study Design
Beckman et al. 2019	On task	Application	Public	SC	ABAB
Callahan et al. 1999	On task	Audio Prompt	NR	GEN	AB
Cihak et al. 2010	Engagement	Picture Prompt	NR	GEN	ABAB & AB
Clemons et al. 2016	On task	Application	Public	GEN	ABAB
Finn et al. 2015	On task	Application	Public	SC	MBD
Holifield et al. 2010	On task	Audio Prompt	Public	SC	MBD
Huffman et al. 2019	On task	Application	Public	GEN	ABAB
Legge et al. 2010	On task	Application	Public	SC	MBD
Romans et al. 2020	On task	Application	Public	SC	ABAB
Rosenbloom et al. 2017	On task	Application	Private	SC	ABAB
Rosenbloom et al. 2016	On task	Application	Public	GEN	ABAB
Siko, 2018	On task	Application	Public	GEN	MBD
Stasolla et al. 2014	On task	Audio Prompt	NR	GEN	MBD
Xu et al. 2017	Engagement	Gestural Prompt	Public	GEN	AB

Note. GEN stands for general education classroom; SC stands for self-contained; NR stands for not reported. For study designs MBD is Multiple Baseline designs and ABAB are reversal designs.

Table 2

Summary of Included Participants

Study/Case	Age	Gender	Race/ Ethnicity	Diagnosis	SM Subject	Student Self-report Time (seconds)
Beckman/Brian	10	Male	White	ASD	LA	15
Beckman/Cody	11	Male	White	FXS; ASD	MAT	20
Callahan/Seth	8	Male	NR	ASD	LA; MAT	15-120 ($M = 60$)
Cihak/Adam	11	NR	NR	HFA	LA; MAT	30
Cihak/Jordan	11	NR	NR	HFA	LA; SS; MAT	30
Cihak/Richard	13	NR	NR	HFA	LA; SC; MAT	30
Clemons/Brad	17	Male	NR	ASD	LA	120
Finn/Adam	8	Male	White	ASD	LA	120
Finn/Bill	8	Male	White	ASD	LA	120
Finn/Paul	9	Male	Hispanic	ASD	LA	120
Finn/Tom	8	Male	Hispanic	ASD	LA	120
Holifield/Graham	9	Male	NR	ASD	LA; MAT	300
Holifield/Tony	10	Male	NR	ASD	LA; MAT	300
Huffman/Simon	19	Male	NR	ASD	NR	60
Legge/Adam	13	Male	NR	ASD	MAT	120
Legge/Joshua	11	Male	NR	ASD	MAT	120
Romans/Jacob	17	Male	NR	PDD-NOS; ADHD; ID	MAT	30
Romans/Zane	15	Male	NR	ASD; ADHD	LA	30
Rosenbloom2017/Carl	17	Male	NR	ASD	LA	30
Rosenbloom2017/Colin	13	Male	NR	ASD	LA	30
Rosenbloom2017/Jack	11	Male	NR	ASD	LA	30
Rosenbloom2017/Stan	10	Male	NR	ASD	LA	30
Rosenbloom2016/Carl	17	Male	African American	ASD	LA	30
Siko/Bobby	22	Male	NR	ASD	SC	360
Siko/Dillon	19	Male	NR	ASD	LA	360
Siko/Seth	20	Male	NR	ASD	COM	360
Stasolla/Jack	7	Male	NR	ASD; ADHD	NR	Varying ($M = 10$)
Stasolla/Michael	8	Male	NR	ASD; ADHD	NR	Varying ($M = 10$)
Xu/Shawn	9	Male	Asian	ASD	LA	60

Note. HFA stands for high functioning autism; PDD-NOS stands for pervasive developmental disorder not otherwise specified; ADHD stands for attention deficit hyperactivity disorder; FXA stands for Fragile X syndrome; LA stands for Language Arts; MAT stands for Math; SS stands for social studies; SC stands for science; COM stands for communication

Risk of Bias across and within included studies are reported in Figures 2 and 3. Across studies, the risk of bias is generally low for data sampling, dependent variable reliability, and selective outcome reporting domains. An unclear risk of bias occurred primarily for sequence generation and participation selection domains. Most studies had a high risk of bias in the blinding of outcome assessors and the blinding of participants and personnel.

Figure 2
Risk of Bias Amongst Studies

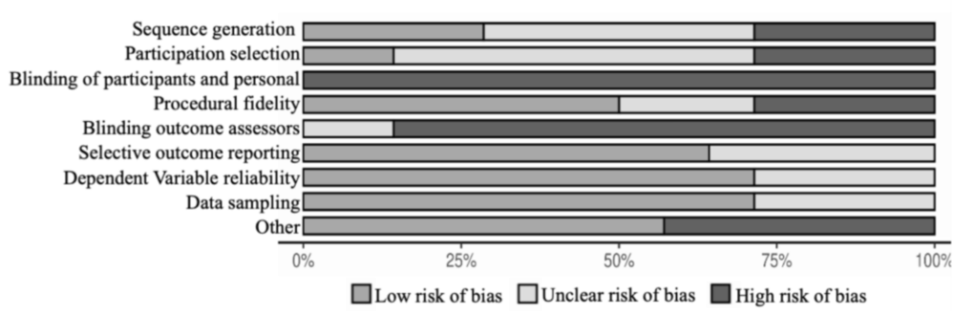
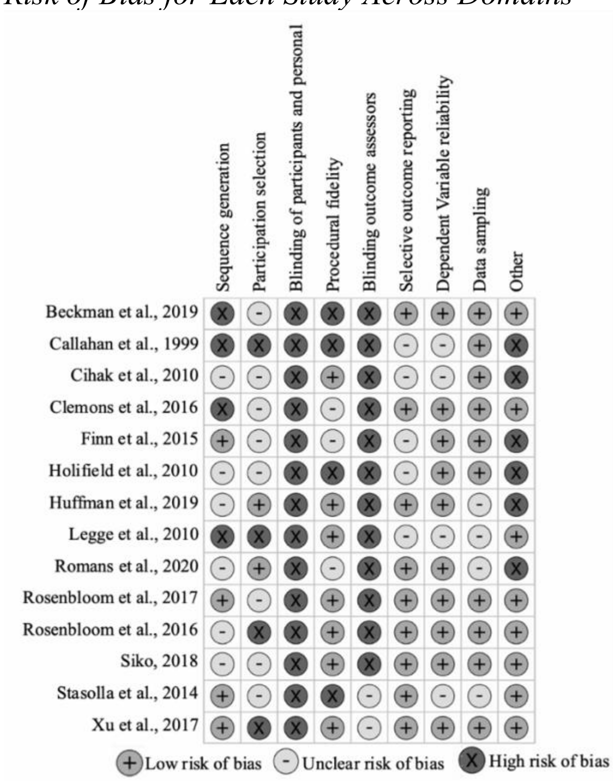


Figure 3
Risk of Bias for Each Study Across Domains



The parameter estimates for academic behavior are presented in Table 3. When modeling the changes of level between the baseline and intervention phase (Model 1), the overall baseline level (γ_{000}) for academic behavior is 37.11% (SE = 4.26, 95% CI [28.16, 45.65], $p < .001$). The overall increase in performance between the baseline and intervention level (treatment effect; γ_{100}) was 49.87% (SE = 4.33, 95% CI [39.76, 59.59], $p < .001$). Although the SM intervention, in general, is large and statistically significant, there is a good deal of between-case and between-study variance in both baseline performance ($\hat{\sigma}^2_{u_0}$ = 18.20, 95% CI [14.13, 22.41]; $\hat{\sigma}^2_{v_0}$= 9.55, 95% CI [3.50, 14.17]) and intervention effectiveness ($\hat{\sigma}^2_{u_1}$= 9.52, 95% CI [6.20, 13.49]; $\hat{\sigma}^2_{v_1}$= 16.57, 95% CI [11.81, 12.91]). This indicates that individual participants and studies differed in their baseline levels of academic behavior and that there was large variability in the treatment effect at the individual participant and study levels.

Table 3

Summary of Three-Level Meta-Analysis

	Model 1		Model 2	
	Estimate (SE)	t	Estimate (SE)	t
Fixed Effect				
Intercept, γ_{000}	37.11(4.26)***	8.46	37.42(5.06)***	7.40
Treatment, γ_{100}	49.87(4.33)***	10.06	55.38(10.94)***	5.06
Random Effects	**Estimate**		**Estimate**	
Between-Cases				
Intercept, $\hat{\sigma}^2_{u_0}$	18.20		17.21	
Treatment, $\hat{\sigma}^2_{u_1}$	9.52		12.91	
Between-Study				
Intercept, $\hat{\sigma}^2_{v_0}$	9.55		7.30	
Treatment, $\hat{\sigma}^2_{v_1}$	16.57		13.90	
Residual Variance, σ^2_e	13.32		13.87	
Moderators				
Classroom Type X Intervention			9.70 (8.68)	
Subject X Intervention			-0.99 (2.28)	
Modality X Intervention			-12.18 (9.37)	
Report Time X Intervention			0.06 (0.07)	

Note. *indicates p <.05, * indicates p <.01, ***indicates p <.001

Given the heterogeneity in the baseline level and treatment effect at the study and case levels, moderators are included in the analysis (Model 2). The four moderators: (1) academic subject during SM intervention (language arts/social studies/communication or Math/Science), (2) classroom type (special or general education), (3) modality of self-monitoring (i.e., application or other method), and (4) average time interval of students self-monitoring academic behavior in seconds, were considered. Tables 1 and 2 display the included moderators' case or study-level descriptives. Although communication is not inherently an academic subject, it was included for one study because the student was participating in an SM intervention during a college-level academic course on communication (Siko, 2018).

The time interval in seconds of student self-monitoring was mean-centered, ranging between -89.21 and 255.93 (range = 345.13). The skewness and kurtosis equal 0.41 and 1.35, respectively, indicating no evidence for deviations from normality. A histogram plot displayed two outlying studies (Finn et al., 2015; Legge et al., 2010). The mean centered frequency for the two participants from Holifield et al. (2010) was 199.7931, and the mean centered frequency for the three cases from Siko et al. (2018) was 259.7931. Therefore, the data from these cases are not considered for meta-regression as they deviate significantly from the overall distribution. Additionally, three studies were excluded from the analysis due to missing values for the moderator's implementation subject and class types (Huffman et al., 2019; Stasolla et al., 2014; Xu et al., 2017).

A total of 9 studies with 17 study participants remained for the Model 2 meta-regression. A reference group was created. The reference group consisted of participants who were in a general education classroom, participated in the intervention during a non-STEM subject, used an application-based SM reporting, and reported their academic behavior at the meantime interval for student self-reporting. For this reference group, the estimated baseline level is 37.42% (SE = 16.31, $t = 2.433$, $p = .015$). The intervention effect for the reference group is 55.38% (SE = 10.94, $t = 5.062$, $p < .001$). The intervention effect is 9.7% higher for students who took part in the SM intervention in a special education classroom compared to a general education classroom, 0.99% lower for students who participated in SM during STEM vs. not, and 12.18% lower for participants who used a non-application-based reporting method (e.g., paper checklist). Lastly, for every unit increase in the self-reporting time interval, the intervention effect increased by 0.06%. None of the moderators are statistically significant. With limited studies and participants, Model 2 analysis had low statistical power. Therefore, consideration of the practical significance was necessary. Of the four moderators across the implementation of SM, the moderator effects are relatively small and do not explain much variability, meaning that these implementation factors do not hold much weight in the effectiveness of SM for autistic students' academic behavior.

Discussion

A multilevel meta-analysis was performed to investigate the effectiveness of SM on the academic engagement of autistic students. Based on previous findings, we expected that SM would positively impact students' academic engagement (Carr et al., 2014; Chia et al., 2018; Scheibel et al., 2024). Overall, results supported this hypothesis, indicating that SM has a strong positive effect on academic behavior (i.e., on-task and engagement) for autistic students.

Corroborating findings in other reviews, we also found that on-task behaviors were the most commonly used outcome measures for SM (Briesch et al., 2018; Scheibel et al., 2024). As on-task behavior is a component of engagement (Fredricks, 2022) and studies operationalized on-task behaviors and academic engagement in the same way, we therefore combined the two studies in which authors reported measuring academic engagement (Cihak et al. 2010; Xu et al. 2017) with the remaining twelve studies that were reported as measuring on-task behavior (Beckman et al., 2019; Callahan et al., 1999; Clemons et al., 2016; Finn et al., 2015; Holifield et al., 2010; Huffman et al., 2019; Legge et al., 2010; Romans et al., 2020; Rosenbloom et al., 2017; Rosenbloom et al., 2016; Siko, 2018; Stasolla et al., 2014). On task behavior and engagement across these studies has included (but is not limited to) sitting in a seat, making eye contact, writing on a worksheet, reading assignment, looking at work, counting manipulatives, and listening to their teacher. Specific behaviors measured for engagement and on-task behavior for each study are reported on OSF.

Autistic students displayed varying academic behavior initially amongst participants and between studies, further corroborating the heterogeneity of academic profiles of autistic students (e.g., Chen et al., 2018; Kim et al., 2018). Considerable variability of the intervention effect on academic behavior was also observed. Although SM was predominantly effective at increasing academic behavior across studies and individuals, it may not be effective for all autistic students. As with all evidence-based practices, including SM, it becomes necessary for practitioners, including teachers, to consider students' unique profiles when selecting interventions (Wong et al., 2015). Future work can investigate learners' characteristics to better inform whom SM strategies may best support.

Extending knowledge from previous meta-analyses (Carr et al., 2014; Chia et al., 2018; Scheibel et al., 2024), we quantitatively investigated some of the characteristics of the intervention’s implementation. Specifically, to gain a broader understanding of how factors in SM implementation may play a role in the intervention’s effectiveness. We anticipated that classroom type (general or special education) and classroom subject (math/science or language arts/social studies/communication) would not impact the effectiveness of SM interventions. Students who partook in SM in a special education classroom had a slightly larger intervention effect than students in a general education classroom. There was also a minimal difference in the increase in academic behavior between subjects. For the most part, the slight differences in effectiveness between classroom settings and subjects further emphasize the effectiveness of SM across environments (Carr et al., 2014; Rafferty & Asaro-Saddler, 2021). Therefore, SM strategies have the potential to support autistic students across their school day, and the variety of classroom settings students may learn within. With the positive effect across settings, there is a need to address the limited use of SM in classroom settings (Schulze, 2016; Southall & Gast, 2011). This can be done by incorporating more learning on evidence-based practices such as SM in teacher education programs (Scheeler et al., 2016). Additionally, it may be worth investigating professional development opportunities for teachers and their effectiveness in promoting the use of SM for current teachers.

Interestingly, we also discovered that application compared to non-application student self-monitoring accounted for little variability in the effectiveness of SM. Across studies, researchers employed a variety of application-based SM tools. The most frequently used was I-Connect (n = 6; Beckman et al., 2019; Clemons et al., 2016; Huffman et al., 2019; Romans et al., 2020; Rosenbloom et al., 2017; Rosenbloom et al., 2016). Other applications included the MotivAider (Legge et al., 2010), the WatchMinder (Finn et al., 2015), and a custom watch application developed at George Mason University for the Android Wear Smartwatch 3 (Siko, 2018). This suggests that a range of modalities - such as applications, pen and paper checklists, and prompts – may support students' self-monitoring.

Additionally, time intervals for self-monitoring during tasks varied widely across studies, with frequency having little impact on effectiveness. Therefore, teachers may consider using a variety of time intervals and self-monitoring methods when implementing SM. Moreover, in addition to teachers using a variety of time intervals, students may also select their own monitoring durations. Allowing students to choose the length of their self-monitoring intervals can enhance the social validity of the intervention while promoting autonomy and self-determination (Wehmeyer et al., 2010).

It is important to note that in our meta-analysis much of the variability in SM intervention effect was not explained by the implementation factors included in our meta-analysis. This suggests that other influences, such as individual characteristics of the learner or classroom, may better account for the variability of SM effectiveness (Scheibel et al., 2024). Although these factors were not explored in the current study, future research may consider examining variables such as classroom size, students' socioeconomic backgrounds, and cognitive developmental profiles. Identifying such moderators can support teachers in making informed decisions when selecting and adapting SM practices to support individual learners.

Overall, SM can be an ideal intervention for autistic students in classroom settings as it provides teachers with lots of flexibility to consider the unique profiles of students and classroom functioning when making implementation-specific decisions, such as classroom type, learning subject, self-monitoring method, and the time interval of students self-reporting.

Limitations

Although this study contributed to the field by clarifying the adaptability and individualized benefit of SM across subjects and classroom types, it has limitations. First, as we conducted a meta-analysis, we needed to depend on information provided in published studies. As three studies were missing implementation subject or class-type intervention information, we excluded these three studies from our moderator analysis. The limited number of studies and study participants with moderators may have led to an underpowered analysis and yielded non-statistically significant moderator effects (Moeyaert et al., 2014). Moreover, even with a large-scale meta-analysis, moderator effects as small as those found in the current study likely would not become statistically significant (Moeyaert et al., 2023).

The studies in this meta-analysis have various ways of defining academic behavior. Outcome measures spanning many components can make it difficult to parse the intervention's effect on specific academic behaviors across studies. This results in a largely general understanding of the effect of SM on academic engagement.

To enhance the interpretability of categorical variables, classroom type, subject, and self-monitoring modality were dummy-coded. This limits the nuanced understanding of the influence of these essential characteristics. However, these moderators still provide a preliminary understanding of the lack of variability among key attributes of the intervention implementation. In our investigation of moderators, we were not able to explain intervention heterogeneity between studies and between cases using the current moderators. Other moderators would be needed to do so. However, it is hard to do so if the information is not reported in the primary studies. It is, therefore, essential that studies report characteristics to allow for future investigations of factors that explain the variability of the effectiveness of SM.

Additionally, there was a high risk of bias in the reporting of outcomes, as those recording behavior were often involved in the training and implementation of the intervention, making them aware of the study phase during data collection. This awareness can introduce bias into the data collection process. Future studies should aim to use procedures in which those reporting observed behaviors are not privy to the studies phase (i.e., baseline or intervention) to reduce potential bias. However, masking participants from phases may not always be possible—particularly when materials specific to the intervention phase (e.g., checklists, tablet applications) are not used during the baseline phase, making the conditions easily distinguishable. Furthermore, in SM studies participant masking is inherently impossible, as participants are explicitly taught to observe and track behaviors that are also being measured by the researcher; therefore, there is an overall high risk of bias in reporting of SM outcomes.

Conclusion

This multilevel meta-analysis allowed for an understanding of the effectiveness of SM for autistic students while using advances in methodology that consider the nested nature of single-case design studies. Overall, SM enhances autistic students' academic behavior (i.e., on-task and engagement) in classroom settings. Moreover, multilevel meta-analysis supported the testing of moderators to explain variability in effectiveness. We found that students' academic behavior improved across classroom type (self-contained or general education), academic subject, modality of self-reporting (i.e., application or other methods), and rate of the time interval of self-reporting, empirically supporting that SM is an effective intervention practices that can be adapted to support individual learner's needs.

References

*References with an asterisk denote studies that were included in the current meta-analysis.

Adams, D., Stainer, M. J., Simpson, K., Paynter, J., & Westerveld, M. F. (2023). Profiles and longitudinal growth trajectories of teacher-rated academic skills and enablers in autistic children and adolescents. *Journal of Autism and Developmental Disorders, 55*(1), 267-283. https://doi.org/10.1007/s10803-023-06186-1

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

Bangert-Drowns, R. L., & Pyke, C. (2001). A taxonomy of student engagement with educational software: An exploration of literate thinking with electronic text. *Journal of Educational Computing Research*, *24*(3), 213–234. https://doi.org/10.2190/0CKM-FKTR-0CPF-JLGR

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.i01.

*Beckman, A. K., Mason, B. A., Wills, H. P., Garrison-Kane, L., & Huffman, J. M. (2019). Improving behavioral and academic outcomes for students with autism spectrum disorder: Testing an app-based self-monitoring intervention. *Education and Treatment of Children, 42*(2), 225-244. https://www.jstor.org/stable/26623055

Bohlmann, N. L., & Downer, J. T. (2015). Self-Regulation and Task Engagement as Predictors of Emergent Language and Literacy Skills. *Early Education and Development, 27*(1), 18–37. https://doi.org/10.1080/10409289.2015.1046784

Briesch, A. M., Daniels, B., & Beneville, M. (2018). Unpacking the term "self-management":

Understanding intervention applications within the school-based literature. *Journal of Behavioral Education, 28*(1), 54-77. https://doi.org/10.1007/s10864-018-9303-1

*Callahan, K. and Rademacher, J. (1999). Using self-management strategies to increase the on-task behavior of a student with autism. *Journal of Positive Behavior Interventions, 1*(2), 117-122. https://doi.org/10.1177/109830079900100206

Carnahan, C., Musti-Rao, S., & Bailey, J. (2008). Promoting active engagement in small group learning experiences for students with autism and significant learning needs. *Education and Treatment of Children, 32*(1), 37-61. https://doi.org/10.1353/etc.0.0047

Carr, M. E., Moore, D. W., & Anderson, A. (2014). Self-management interventions on students with autism: A meta-analysis of single-subject research. *Exceptional Children*, 81(1), 28-44. https://doi.org/10.1177/0014402914532235

Chen, L., Abrams, D. A., Rosenberg-Lee, M., Iuculano, T., Wakeman, H. N., Prathap, S., Chen, T., & Menon, V. (2019). Quantitative analysis of heterogeneity in academic achievement of children with autism. *Clinical Psychological Science, 7*(2), 362–380. https://doi.org/10.1177/2167702618809353

Chia, G. L. C., Anderson, A., & McLean, L. A. (2018). Use of technology to support self-management in individuals with autism: systematic review. *Review Journal of Autism and Developmental Disorders, 5*(2), 142-155. https://doi.org/10.1007/s40489-018-0129-5

*Cihak, D. F., Wright, R., & Ayres, K. M. (2010). Use of Self-Modeling Static-Picture Prompts via a Handheld Computer to Facilitate Self-Monitoring in the General Education Classroom. *Education and Training in Autism and Developmental Disabilities*, *45*(1), 136–149. http://www.jstor.org/stable/23880157

*Clemons, L. L., Mason, B. A., Garrison-Kane, L., & Wills, H. P. (2015). Self-monitoring for high school students with disabilities. *Journal of Positive Behavior Interventions*, *18*(3), 145-155. https://doi.org/10.1177/1098300715596134

Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis* (3rd ed.). Pearson.

Estes, A., Rivera, V., Bryan, M., Cali, P. W., & Dawson, G. (2010). Discrepancies between academic achievement and intellectual ability in higher-functioning school-aged children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 41*(8), 1044-1052. https://doi.org/10.1007/s10803-010-1127-3

*Finn, L., Ramasamy, R., Dukes, C., & Scott, J. T. (2014). Using watchminder to increase the on-task behavior of students with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(5), 1408-1418. https://doi.org/10.1007/s10803-014-2300-x

Fredricks, J.A. (2022). The measurement of student engagement: Methodological advances and comparison of new self-report instruments. In: Reschly, A.L., Christenson, S.L. (eds) *Handbook of Research on Student Engagement.* Springer, Cham. https://doi.org/10.1007/978-3-031-07853-8_29

*Holifield, C., Goodman, J. C., Hazelkorn, M., & Heflin, L. J. (2010). Using self-monitoring to increase attending to task and academic accuracy in children with autism. *Focus on Autism and Other Developmental Disabilities, 25*(4), 230-238. https://doi.org/10.1177/1088357610380137

*Huffman, J. M., Bross, L. A., Watson, E. K., Wills, H. P., & Mason, R. A. (2019). Preliminary investigation of a self-monitoring application for a postsecondary student with autism. *Advances in Neurodevelopmental Disorders, 3*(4), 423-433. https://doi.org/10.1007/s41252-019-00124-y

Hume, K., Steinbrenner, J. R., Odom, S. L., Morin, K. L., Nowell, S. W., Tomaszewski, B., Szendrey, S., McIntyre, N. S., Yücesoy-Özkan, S., & Savage, M. N. (2021). Evidence-based practices for children, youth, and young adults with autism: Third generation review. *Journal of Autism and Developmental Disorders, 51*(11), 4013–4032. https://doi.org/10.1007/s10803-020-04844-2

Keen, D., Adams, D., & Simpson, K. (2023). Teacher ratings of academic skills and academic enablers of children on the autism spectrum. *International Journal of Inclusive Education, 27*(10), 1085-1101. https://doi.org/10.1080/13603116.2021.1881626

Kim, S. H., Bal, V. H., & Lord, C. (2018). Longitudinal follow-up of academic achievement in children with autism from age 2 to 18. *Journal of Child Psychology and Psychiatry, 59*(3), 258-267. https://doi.org/10.1111/jcpp.12808

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Lawson, M. A. & Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. *Review of Educational Research, 83*(3), 432-479. https://doi.org/10.3102/0034654313480891

Lawson, M. A. & Masyn, K. E. (2015). Analyzing profiles, predictors, and consequences of student engagement dispositions. *Journal of School Psychology, 53*(1), 63-86. https://doi.org/10.1016/j.jsp.2014.11.004

Ledford, J. R., Hall, E., Conder, E., & Lane, J. D. (2016). Research for young children with autism spectrum disorders: Evidence of social and ecological validity. *Topics in Early Childhood Special Education, 35*(4), 223-233.

https://doi.org/10.1177/0271121415585956

Lee, S. H., Simpson, R. L., & Shogren, K. A. (2007). Effects and implications of self-management for students with autism: A meta-analysis. *Focus on Autism and Other Developmental Disabilities*, *22*, 2–13. https://doi.org/10.1177/10883576070220010101

*Legge, D. B., DeBar, R. M., & Alber-Morgan, S. R. (2010). The effects of self-monitoring with a motivaider® on the on-task behavior of fifth and sixth graders with autism and other disabilities. *Journal of Behavior Assessment and Intervention in Children, 1*(1), 43-52. https://doi.org/10.1037/h0100359

Lenz, A. S. (2013). Calculating effect size in single-case research. *Measurement and Evaluation in Counseling and Development, 46*(1), 64-73. https://doi.org/10.1177/0748175612456401

Lin, L. I.-K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, *45*(1), 255–268. https://doi.org/10.2307/2532051

Mayes, S. D. & Calhoun, S. L. (2007). Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. *Child Neuropsychology, 13*(6), 469-493. https://doi.org/10.1080/09297040601112773

McDougall, D. (1998). Research on self-management techniques used by students with disabilities in general education settings: A descriptive review. *Remedial and Special Education, 19*(5), 310-320. https://doi.org/10.1177/074193259801900507

McGuinness, L. A. & Higgins, J. P. T. (2020). Risk-of-bias visualization (robvis): An r package and shiny web app for visualizing risk-of-bias assessments. *Research Synthesis Methods, 12*(1), 55-61. https://doi.org/10.1002/jrsm.1411

Moeyaert, M., Ferron, J. M., Beretvas, S. N., & Van den Noortgate, W. (2014). From a single-level analysis to a multilevel analysis of single-case experimental designs. *Journal of School Psychology*, *52*(2), 191–211. https://doi.org/10.1016/j.jsp.2013.11.003

Moeyaert, M., Maggin, D., & Verkuilen, J. (2016). Reliability, validity, and usability of data extraction programs for single-case research designs. *Behavior Modification*, *40*(6), 874-900. https://doi.org/10.1177/0145445516645763

Moeyaert, M., Yang, P., & Xue, Y. (2023). Individual participant data meta-analysis including moderators: Empirical validation. *The Journal of Experimental Education, 92*(4), 723-740. https://doi.org/10.1080/00220973.2023.2208062

Moeyaert, M., Zimmerman, K. N., & Ledford, J. R. (2018). *Synthesis and meta-analysis of single case research*. In J. R. Ledford & D. L. Gast (Eds), Single case research methodology, (pp. 393-416). Routledge. https://doi.org/10.4324/9781315150666

Morningstar, M. E., Lombardi, A., Fowler, C. H., & Test, D. W. (2017). A college and career readiness framework for secondary students with disabilities. *Career Development and Transition for Exceptional Individuals, 40*(2), 79-91.

National Autism Center. (2015). Findings and conclusions: National standards project, Phase 2. Randolph, MA: Author

National Center for Education Statistics. (2023). Students with Disabilities. *Condition of Education.* U.S. Department of Education, Institute of Education Sciences. Retrieved June 25, 2023, from https://nces.ed.gov/programs/coe/indicator/cgg

Oberman, L. M., & Kaufmann, W. E. (2020). Autism spectrum disorder versus autism spectrum disorders: Terminology, concepts, and clinical practice. *Frontiers in Psychiatry*, *11*(484), 1-5. https://doi.org/10.3389/fpsyt.2020.00484

Odom, S. L., Hume, K., Boyd, B. A., & Stabel, A. (2012). Moving beyond the intensive behavior treatment versus eclectic dichotomy. *Behavior Modification, 36*(3), 270-297. https://doi.org/10.1177/0145445512444595

OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]. https://chat.openai.com/chat

Ouzzani, M., Hammady, H. M., Fedorowicz, Z., & Elmagarmid, A. K. (2016). Rayyan—a web and mobile app for systematic reviews. Systematic Reviews, 5(1). https://doi.org/10.1186/s13643-016-0384-4

Pustejovsky, J. E. (2018). Using response ratios for meta-analyzing single-case designs with behavioral outcomes. *Journal of School Psychology, 68*, 99-112. https://doi.org/10.1016/j.jsp.2018.02.003

Pustejovsky, J. E. (2019). Procedural sensitivities of effect sizes for single-case designs with directly observed behavioral outcome measures. *Psychological Methods, 24*(2), 217–235. https://doi.org/10.1037/met0000179

Rafferty, L. A., & Asaro-Saddler, K. (2021). Effective self-management strategies. In U. Sharma & L. Maheady (Eds.), *Oxford Research Encyclopedia of Inclusive and Special Education.* New York: Oxford University Press.

Rakap, S. (2015). Effect sizes as result interpretation aids in single-subject experimental research: description and application of four nonoverlap methods. *British Journal of Special Education*, *42*(1), 11-33. https://doi.org/10.1111/1467-8578.12091

Reichow, B., Barton, E. E., & Maggin, D. M. (2018). Development and applications of the single-case design risk of bias tool for evaluating single-case design research study reports. *Research in Developmental Disabilities, 79*, 53-64.

https://doi.org/10.1016/j.ridd.2018.05.008

Rohatgi (2022). WebPlotDigitizer 4.6 [Computer software manual]. Available from https://automeris.io/WebPlotDigitizer

* Romans, S. K., Wills, H. P., Huffman, J. M., & Garrison-Kane, L. (2020). The effect of web-based self-monitoring to increase on-task behavior and academic accuracy of high school students with autism. *Preventing School Failure: Alternative Education for Children and Youth*, 64(3), 249–260. https://doi.org/10.1080/1045988X.2020.1732282

*Rosenbloom, R. (2017). *The effects of a technology-based self-monitoring intervention on on-task, disruptive, and task-completion behaviors for adolescents with autism.* (publication No. 2013956707) [Doctoral Dissertation, University of Kansas]. ProQuest Dissertations & Theses Global.

* 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

Scheeler, M. C., Budin, S., & Markelz, A. (2016). The role of teacher preparation in promoting evidence-based practice in schools. *Learning Disabilities: A Contemporary Journal*, 14(2), 171-187.

Scheibel, G., Zaeske, L. M., Malone, E. J., & Zimmerman, K. N. (2024). A meta-analysis of self-management interventions for students with ASD. *Research in Autism Spectrum Disorders, 110*, 102294. https://doi.org/10.1016/j.rasd.2023.102294

Schulze, M. A. (2016). Self-management strategies to support students with ASD. *Teaching Exceptional Children, 48*(5), 225-231. https://doi.org/10.1177/0040059916640759

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education*, *8*(2), 24-33. https://doi.org/10.1177/074193258700800206

*Siko, L. (2018). *Wearable technology for self-monitoring of academic engagement in college students with autism spectrum disorder* (Publication No. 10814889) [Doctoral dissertation, George Mason University]. ProQuest Dissertations & Theses Global.

Southall, C. M., & Gast, D. L. (2011). Self-management procedures: A comparison across the autism spectrum. *Education and Training in Autism and Developmental Disabilities*, *46*(2), 155-171. https://www.jstor.org/stable/23879688

Sparapani, N., Morgan, L., Reinhardt, V. P., Schatschneider, C., & Wetherby, A. M. (2016). Evaluation of classroom active engagement in elementary students with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*, 782-796. https://doi.org/10.1007/s10803-015-2615-2

*Stasolla, F., Perilli, V., & Damiani, R. (2014). Self monitoring to promote on-task behavior by two high functioning boys with autism spectrum disorders and symptoms of ADHD. *Research in Autism Spectrum Disorders, 8*(5), 472-479. https://doi.org/10.1016/j.rasd.2014.01.007

Steinbrenner, J. R. and Watson, L. R. (2015). Student engagement in the classroom: The impact of classroom, teacher, and student factors. *Journal of Autism and Developmental Disorders, 45*(8), 2392-2410. https://doi.org/10.1007/s10803-015-2406-9

Steinbrenner, J. R., Hume, K., Odom, S. L., Morin, K. L., Nowell, S. W., Tomaszewski, B., Szendrey, S., McIntyre, N. S., Yücesoy-Özkan, S., & Savage, M. N. (2020). Evidence-based practices for children, youth, and young adults with Autism. The University of

North Carolina at Chapel Hill, Frank Porter Graham Child Development Institute, National Clearinghouse on Autism Evidence and Practice Review Team.

Wang, S., Scells, H., Koopman, B., & Zuccon, G. (2023). Can ChatGPT write a good boolean query for systematic review literature search?. https://doi.org/10.1145/3539618.3591703

Wehmeyer, M. L., Shogren, K. A., Zager, D., Smith, T. E., & Simpson, R. (2010). Research-based principles and practices for educating students with autism: Self-determination and social interactions. *Education and Training in Autism and Developmental Disabilities*, 475-486. https://www.jstor.org/stable/23879754

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., Brock, M. E., Plavnick, J. B., Fleury, V. P., & Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of Autism and Developmental Disorders*, *45*, 1951-1966. https://doi.org/10.1007/s10803-014-2351-z

* Xu, S., Wang, J., Lee, G. T., & Luke, N. (2017). Using self-monitoring with guided goal setting to increase academic engagement for a student with autism in an inclusive classroom in China. *The Journal of Special Education*, *51*(2), 106-114. https://doi.org/10.1177/0022466916679980